Available online at www.sciencedirect.com
SCIENCE DIRECT®
Reliability Engineering and System Safety 81 (2003) 23–69
RELIABILITY ENGINEERING & SYSTEM SAFETY
www.elsevier.com/locate/ress

Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems

J.C. Helton[a,*], F.J. Davis[b]

[a]Department of Mathematics and Statistics, Arizona State University, Tempe, AZ 85287-1804, USA
[b]Department 6849, MS 0779, Sandia National Laboratories, Albuquerque, NM 87185-0779, USA

Received 28 January 2003; accepted 25 February 2003

Abstract

The following techniques for uncertainty and sensitivity analysis are briefly summarized: Monte Carlo analysis, differential analysis, response surface methodology, Fourier amplitude sensitivity test, Sobol' variance decomposition, and fast probability integration. Desirable features of Monte Carlo analysis in conjunction with Latin hypercube sampling are described in discussions of the following topics: (i) properties of random, stratified and Latin hypercube sampling, (ii) comparisons of random and Latin hypercube sampling, (iii) operations involving Latin hypercube sampling (i.e. correlation control, reweighting of samples to incorporate changed distributions, replicated sampling to test reproducibility of results), (iv) uncertainty analysis (i.e. cumulative distribution functions, complementary cumulative distribution functions, box plots), (v) sensitivity analysis (i.e. scatterplots, regression analysis, correlation analysis, rank transformations, searches for nonrandom patterns), and (vi) analyses involving stochastic (i.e. aleatory) and subjective (i.e. epistemic) uncertainty.
Published by Elsevier Science Ltd.

Keywords: Aleatory uncertainty; Epistemic uncertainty; Latin hypercube sampling; Monte Carlo analysis; Random sampling; Sensitivity analysis; Uncertainty analysis

1. Introduction

The assessment and presentation of the effects of uncertainty are now widely recognized as important parts of analyses for complex systems [1–6]. At the simplest level, such analyses can be viewed as the study of functions of the form

$$\mathbf{y} = \mathbf{f}(\mathbf{x}), \tag{1.1}$$

where the function $\mathbf{f}$ represents the model or models under study, $\mathbf{x} = [x_1, x_2, \ldots]$ is a vector of model inputs, and $\mathbf{y} = [y_1, y_2, \ldots]$ is a vector of model predictions. The goal of an uncertainty analysis is to determine the uncertainty in the elements of $\mathbf{y}$ that results from uncertainty in the elements of $\mathbf{x}$. A typical adjunct to an uncertainty analysis is a sensitivity analysis, which attempts to determine how the uncertainty in individual elements of $\mathbf{x}$ affects the uncertainty in the elements of $\mathbf{y}$. In practice, $\mathbf{f}$ can be quite complex (e.g. one or more computer programs involving complex algorithms and many thousands of lines of programming); further, $\mathbf{x}$ and $\mathbf{y}$ are often of high dimension.

To carry out uncertainty and sensitivity analyses, the uncertainty in the elements of $\mathbf{x}$ must be characterized. For this presentation, the uncertainty in the elements of $\mathbf{x}$ is assumed to be characterized by a sequence of distributions

$$D_1, D_2, \ldots, D_{nX}, \tag{1.2}$$

where D_j is the distribution associated with the element x_j of $\mathbf{x}$ and nX is the number of elements contained in $\mathbf{x}$ (i.e. $\mathbf{x} = [x_1, x_2, \ldots, x_{nX}]$). Various correlations and additional relationships between the elements of $\mathbf{x}$ are also possible. Initially, the distributions in Eq. (1.2) will be assumed to characterize a degree of belief with respect to where the appropriate values for the elements of $\mathbf{x}$ are located for use in the evaluation of the function $\mathbf{f}$ in Eq. (1.1). When used in this manner, these distributions are providing a quantitative representation for what is commonly referred to as subjective or epistemic uncertainty [7,8]. Such distributions are often developed through an expert review process [9–29].

* Corresponding author. Address: Department 6849, MS 0779, Sandia National Laboratories, Albuquerque, NM 87185-0779, USA. Tel.: +1-505-284-4808; fax: +1-505-844-2348.
E-mail address: jchelto@sandia.gov (J.C. Helton).

0951-8320/03/$ - see front matter Published by Elsevier Science Ltd.
doi:10.1016/S0951-8320(03)00058-9

For notational convenience, the function **f**, and hence **y**, in Eq. (1.1) will be assumed to be scalar-valued, although such simplicity is almost never the case in real analyses. With this assumption, the representation in Eq. (1.1) becomes

$$y = f(\mathbf{x}). \tag{1.3}$$

Further, again for notational convenience and also for ease in distinguishing between different uses of probability at later points in this presentation, the distributions in Eq. (1.2) and any additional relationships imposed on the elements of **x** will be represented by a probability space $(\mathscr{S}_{su}, \mathfrak{S}_{su}, p_{su})$, where the subscript 'su' is used as a designator for 'subjective'. As a reminder, a probability space $(\mathscr{S}, \mathfrak{S}, p)$, consists of three elements: a set $\mathscr{S}$ that contains everything that could occur in the particular universe under consideration; a collection $\mathfrak{S}$ of subsets of $\mathscr{S}$ for which probability will be defined; and a function p that actually defines probability for the elements of $\mathfrak{S}$ (Sect. IV.4, Ref. [30]). In the terminology of probability theory, the set $\mathscr{S}$ is the sample space; the elements of $\mathscr{S}$ (i.e. the vectors **x** in Eqs. (1.1) and (1.3)) are elementary events; the elements of $\mathfrak{S}$ are events; and the function p is a probability measure.

When viewed in its most general form, uncertainty analysis simply involves determination of the distribution for y that results from the function f in Eq. (1.3) and the distributions $D_1, D_2, \ldots, D_{nX}$ in Eq. (1.2), which define the probability space $(\mathscr{S}_{su}, \mathfrak{S}_{su}, p_{su})$. Further, the distribution for y can be presented as a cumulative distribution function (CDF) or as a complementary cumulative distribution function (CCDF), which is simply one minus the CDF (Fig. 1). The CCDF is typically used when it is desired to display small probabilities associated with large values of y or when it is desired to answer the question "How likely is y to be this large or larger?" Given that it can be determined, the CDF, or equivalently the CCDF, in Fig. 1 provides a complete representation of the uncertainty in y. A density function can also be used to summarize the uncertainty in y; however, CDFs and CCDFs provide more convenient and informative summaries in sampling-based studies.

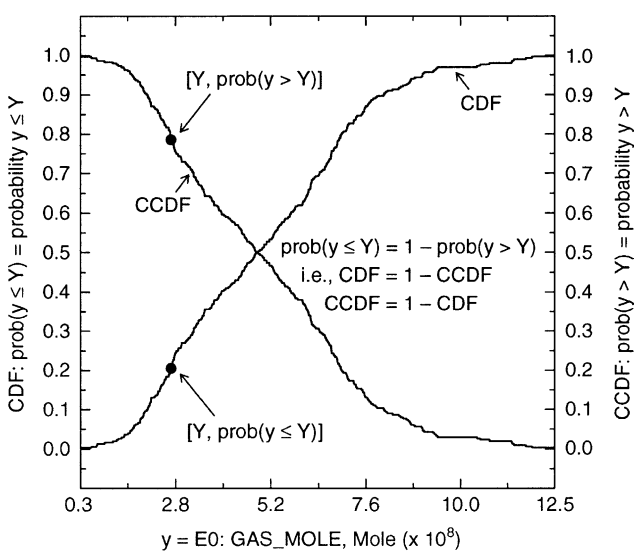

Fig. 1. Use of CDFs and CCDFs to represent uncertainty in model predictions.

The CCDF in Fig. 1 can be formally defined by the integral

$$\text{prob}(y > Y) = \int_{\mathscr{S}_{su}} \delta_Y[f(\mathbf{x})] d_{su}(\mathbf{x}) \mathrm{d}V_{su}, \tag{1.4}$$

where $\text{prob}(y > Y)$ is the probability that a value larger than Y will occur, d_{su} represents the density function corresponding to the distributions in Eq. (1.2) and hence to the probability space $(\mathscr{S}_{su}, \mathfrak{S}_{su}, p_{su})$, the differential $\mathrm{d}V_{su}$ is selected for mnemonic purposes because integration will typically be over a high-dimension (i.e. nX) volume, and

$$\delta_Y[f(\mathbf{x})] = \begin{cases} 1 & \text{if } f(\mathbf{x}) > Y \\ 0 & \text{if } f(\mathbf{x}) \le Y \end{cases}. \tag{1.5}$$

Similarly, the corresponding CDF is defined by

$$\begin{aligned} \text{prob}(y \le Y) &= 1 - \text{prob}(y > Y) \\ &= 1 - \int_{\mathscr{S}_{su}} \delta_Y[f(\mathbf{x})] d_{su}(\mathbf{x}) \mathrm{d}V_{su}, \end{aligned} \tag{1.6}$$

where $\text{prob}(y \le Y)$ is the probability that a value less than or equal to Y will occur. Although the integral in Eqs. (1.4) and (1.6) formally defines the CCDF and CDF associated with y, in practice this integral is not amenable to a closed-form evaluation; rather, some type of approximation procedure must be used. In particular, the focus of this presentation is on the use of Latin hypercube sampling [31,32] in the approximation of this integral.

As just indicated, uncertainty analysis is simple in concept and involves evaluation of the integral in Eq. (1.4) to obtain the CDF and CCDF in Fig. 1. Sensitivity analysis involves the determination of the effects of the individual elements of **x** on $y = f(\mathbf{x})$. Although sensitivity analysis is closely tied to uncertainty analysis, it tends to be a more complex undertaking due to both the variety of possible measures of sensitivity and the additional computational procedures required to evaluate these measures. This presentation will emphasize sensitivity measures that can be obtained when Latin hypercube sampling is used to evaluate the integral in Eq. (1.4).

One formal way to look at sensitivity analysis is to view it as an analysis of variance problem. Specifically, the variance $V(y)$ of y is given by

$$V(y) = \int_{\mathscr{S}_{su}} [E(y) - f(\mathbf{x})]^2 d_{su}(\mathbf{x}) \mathrm{d}V_{su}, \tag{1.7}$$

where $E(y)$ denotes the expected value of y and is given by

$$E(y) = \int_{\mathscr{S}_{su}} f(\mathbf{x}) d_{su}(\mathbf{x}) \mathrm{d}V_{su}. \tag{1.8}$$

Sensitivity analysis can then be viewed as a decomposition of $V(y)$ into components due to the individual elements of $\mathbf{x}$, with the size of these components then providing an indication of variable importance. However, as will be discussed, not all sensitivity analysis procedures are mathematically equivalent to a variance decomposition problem. A variety of sensitivity measures based on Latin hypercube sampling will be presented.

2. Techniques for uncertainty and sensitivity analysis

The use of Latin hypercube sampling constitutes part of what is often called a Monte Carlo procedure for the propagation of uncertainty. In addition, there exist a number of other procedures that are also used for the propagation of uncertainty, including differential analysis, response surface methodology (RSM), the Fourier amplitude sensitivity test (FAST) and the closely related Sobol' variance decomposition, and fast probability integration (FPI). To provide perspective on, and a context for, the use of Latin hypercube sampling, the preceding procedures are briefly summarized in this section.

2.1. Monte Carlo analysis

In Monte Carlo analysis, a probabilistically based sampling procedure is used to develop a mapping from analysis inputs to analysis results. This mapping then provides a basis for both the evaluation of the integral in Eq. (1.4) (i.e. uncertainty analysis) and the evaluation of the effects of individual elements of $\mathbf{x}$ on $y = f(\mathbf{x})$ (i.e. sensitivity analysis). Specifically, a sample

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{i,nX}], \quad i = 1, 2, \ldots, nS, \tag{2.1}$$

of size nS is generated from $\mathcal{S}_{su}$ in consistency with the distributions in Eq. (1.2) (i.e. in consistency with the definition of the probability space $(\mathcal{S}_{su}, \mathfrak{S}_{su}, p_{su})$). A number of possible sampling procedures exist, including random sampling, stratified sampling, and Latin hypercube sampling (see Section 3). The preceding sampling procedures are probabilistically based in the sense that weights

$$w_i, \quad i = 1, 2, \ldots, nS, \tag{2.2}$$

exist such that the result obtained with sample element $\mathbf{x}_i$ can be used in conjunction with the weight w_i to obtain quantities such as expected values, variances and other entities that derive from integration over $\mathcal{S}_{su}$. For random sampling and also Latin hypercube sampling, w_i is the reciprocal of the sample size (i.e. $w_i = 1/nS$); for stratified sampling, w_i is determined by the probability of the stratum (i.e. subset of $\mathcal{S}_{su}$) from which $\mathbf{x}_i$ was sampled and the number of samples taken from that stratum.

Once the sample in Eq. (2.1) is generated, evaluation of f creates the following mapping from analysis inputs to analysis results:

$$[\mathbf{x}_i, y_i], \quad i = 1, 2, \ldots, nS, \tag{2.3}$$

where $y_i = f(\mathbf{x}_i)$. Then, the integrals in Eqs. (1.4), (1.7) and (1.8) can be approximated by

$$\text{prob}(y > Y) \doteq \widehat{\text{prob}}(y > Y) = \sum_{i=1}^{nS} \delta_Y(y_i) w_i \tag{2.4}$$

$$E(y) \doteq \hat{E}(y) = \sum_{i=1}^{nS} y_i w_i \tag{2.5}$$

$$V(y) \doteq \hat{V}(y) = \sum_{i=1}^{nS} [\hat{E}(y) - y_i]^2 w_i. \tag{2.6}$$

The distribution function approximated in Eq. (2.4) provides the most complete representation of the uncertainty in y that derives from the distributions in Eq. (1.2) and hence from the probability space $(\mathcal{S}_{su}, \mathfrak{S}_{su}, p_{su})$. The expected value and variance approximated in Eqs. (2.5) and (2.6) provide a summary of this distribution but with the inevitable loss of resolution that occurs when the information contained in $2nS$ numbers (i.e. in the y_i and w_i) is mapped into two numbers. For random sampling, use of $w_i = 1/(nS - 1)$ in Eq. (2.6) results in an unbiased estimate for $V(y)$.

The mapping in Eq. (2.3) can be explored with various techniques to determine the effects of the individual elements of $\mathbf{x}$ on y. For example, scatterplots based on the points

$$[x_{ij}, y_i], \quad i = 1, 2, \ldots, nS, \tag{2.7}$$

for each element x_j of $\mathbf{x}$ may completely reveal the relationships between the x_j and y (Fig. 2). Another possibility is to use the results in Eq. (2.3) and least squares techniques to construct a regression model of the form

$$y = b_0 + \sum_{j=1}^{nX} b_j x_j \tag{2.8}$$

that relates y to the x_j. Various aspects of this model and the construction process that produced it can then be used to infer the relationships between the x_j and y. The preceding and other techniques for sensitivity analysis in conjunction with Monte Carlo procedures are discussed in more detail in Section 6.

Monte Carlo procedures for the propagation of uncertainty are very popular and many examples of their use exist [33–47]. Further, Monte Carlo procedures find a wide variety of applications in the sciences and a very extensive literature exists [48–60].

2.2. Differential analysis

Differential analysis is based on the partial derivatives of f with respect to the elements of $\mathbf{x}$. In its simplest form, differential analysis involves approximating the model by

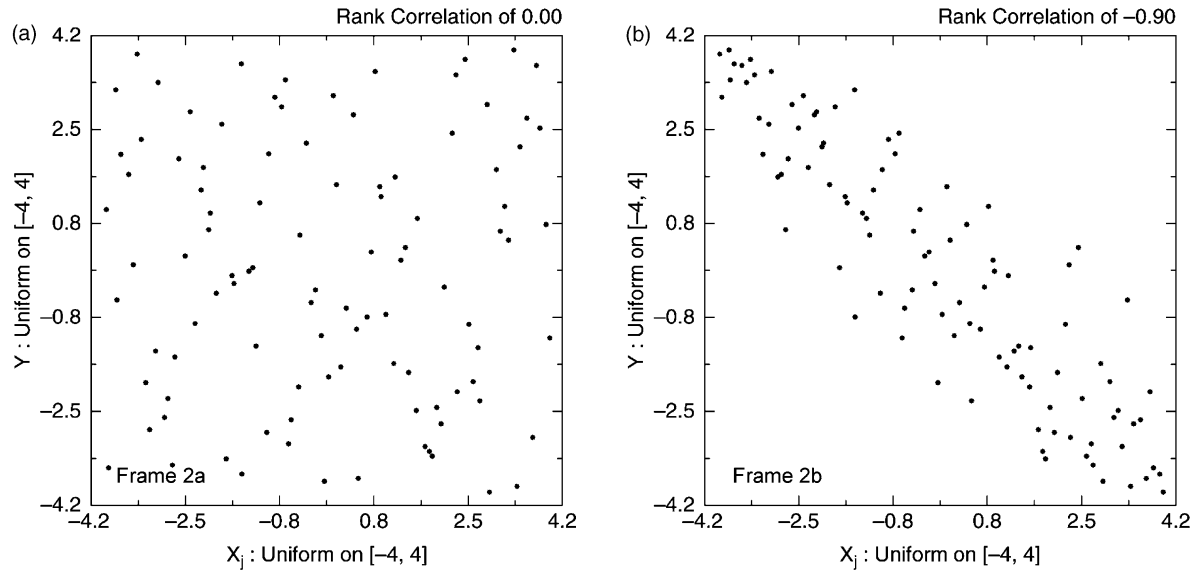

Fig. 2. Scatterplots produced in a Monte Carlo analysis with a Latin hypercube sample of size $nS = 100$: (a) no relationship between x_j and y, and (b) well-defined relationship between x_j and y (see Section 5.1 for a discussion of rank correlation).

the Taylor series

$$y(\mathbf{x}) \doteq f(\mathbf{x}_0) + \sum_{j=1}^{nX} [\partial f(\mathbf{x}_0)/\partial x_j][x_j - x_{j0}], \tag{2.9}$$

where $\mathbf{x}_0 = [\mathbf{x}_{10}, \mathbf{x}_{20}, \ldots, \mathbf{x}_{nX,0}]$ is a vector of base-case values for the x_j (e.g. the expected values for the x_j defined by the distributions in Eq. (1.2)).

Once the approximation in Eq. (2.9) is determined, variance propagation formulas can be used to determine the uncertainty in y that results from the distributions in Eq. (1.2). In particular,

$$E(y) \doteq y(\mathbf{x}_0) + \sum_{j=1}^{nX} [\partial f(\mathbf{x}_0)/\partial x_j] E(x_j - x_{j0}) = y(\mathbf{x}_0) \tag{2.10}$$

and

$$V(y) \doteq \sum_{j=1}^{nX} [\partial f(\mathbf{x}_0)/\partial x_j]^2 V(x_j) + 2 \sum_{j=1}^{nX} \sum_{k=j+1}^{nX} [\partial f(\mathbf{x}_0)/\partial x_j] \times [\partial f(\mathbf{x}_0)/\partial x_k] \mathrm{Cov}(x_j, x_k), \tag{2.11}$$

where E, V and Cov denote expected value, variance and covariance, respectively. If the x_j are uncorrelated, then

$$V(y) \doteq \sum_{j=1}^{nX} [\partial f(\mathbf{x}_0)/\partial x_j]^2 V(x_j). \tag{2.12}$$

Thus, the Taylor series in Eq. (2.9) leads to approximations of the expected value and variance for y that result from the distributions in Eq. (1.2). Differential analysis does not lead very naturally to an approximation for the CDF or CCDF for y, although such approximations could be obtained by using a Monte Carlo simulation of the Taylor series in Eq. (2.9).

The determination of expected values, variances and possibly CDFs or CCDFs constitutes the uncertainty analysis component of differential analysis. Sensitivity analysis is based on the use of the partial derivatives associated with a Taylor series to determine the effects of the individual elements of $\mathbf{x}$ on y. For example, if the Taylor series in Eq. (2.9) is used and the elements of $\mathbf{x}$ are independent, then the fractional contribution of x_j to the variance of y can be approximated by

$$V(y|x_j) \doteq [\partial f(\mathbf{x}_0)/\partial x_j]^2 V(x_j)/V(y), \tag{2.13}$$

with $V(y)$ being obtained from the approximation in Eq. (2.12). An ordering of the x_j on the basis of the size of the fractional contributions $V(y|x_j)$ provides a ranking of variable importance on the basis of how much of the variance of y can be accounted for by each element of $\mathbf{x}$.

Normalization of the partial derivatives in the Taylor series in Eq. (2.9) provides a basis for another approach to assessing the importance of individual elements of $\mathbf{x}$. In particular, the following normalizations are possible:

$$\frac{y(\mathbf{x}) - y(\mathbf{x}_0)}{y(\mathbf{x}_0)} \doteq \sum_{j=1}^{nX} \left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j} \frac{x_{j0}}{y(\mathbf{x}_0)} \right] \left[\frac{x_j - x_{j0}}{x_{j0}} \right] \tag{2.14}$$

and

$$\frac{y(\mathbf{x}) - y(\mathbf{x}_0)}{\mathrm{SD}(y)} \doteq \sum_{j=1}^{nX} \left[\frac{\partial f(\mathbf{x}_0)}{\partial x_j} \frac{\mathrm{SD}(x_j)}{\mathrm{SD}(y)} \right] \left[\frac{x_j - x_{j0}}{\mathrm{SD}(x_j)} \right], \tag{2.15}$$

where SD denotes standard deviation, SD(y) is estimated from Eq. (2.12), and no problem with respect to division by zero exists. The normalized coefficients

$$C_{\mathrm{bc}}(x_j) = [\partial f(\mathbf{x}_0)/\partial x_j][x_{j0}/y(\mathbf{x}_0)], \quad j = 1, 2, \ldots, nX, \tag{2.16}$$

from Eq. (2.14) provide a ranking of variable importance based on equal fractional changes from base-case values x_{j0} and thus incorporate no distributional information about the elements of $\mathbf{x}$. The normalized coefficients

$$C_{sd}(x_j) = [\partial f(\mathbf{x}_0)/\partial x_j][\mathrm{SD}(x_j)/\mathrm{SD}(y)], \quad j = 1, 2, \ldots, nX, \tag{2.17}$$

from Eq. (2.15) provide a ranking of variable importance based on changes from base-case values x_{j0} that are equal fractions of the standard deviation $\mathrm{SD}(x_j)$ of x_j. Thus, unlike rankings of variable importance with the coefficients in Eq. (2.16), rankings with the coefficients in Eq. (2.17) incorporate the distributional assumptions for the elements of $\mathbf{x}$.

The quality of results obtained in a differential analysis is limited by the quality of the underlying Taylor series approximation. In particular, if y is a nonlinear function of the elements of $\mathbf{x}$, then the first-order Taylor series approximation in Eq. (2.9) may provide a poor representation of the relationships between y and the elements of $\mathbf{x}$. Better approximations to y can be obtained by using higher-order Taylor series. For example, a second-order approximation has the form

$$y(\mathbf{x}) \doteq y(\mathbf{x}_0) + \sum_{j=1}^{nX} [\partial f(\mathbf{x}_0)/\partial x_j][x_j - x_{j0}] + \frac{1}{2}\sum_{j=1}^{nX}\sum_{k=1}^{nX} [\partial^2 f(\mathbf{x}_0)/\partial x_j \partial x_k][(x_j - x_{j0})(x_k - x_{k0})]. \tag{2.18}$$

If the preceding approximation to $y(\mathbf{x})$ is used, the elements of $\mathbf{x}$ are uncorrelated, and fourth-order and higher-order terms are ignored in the derivation of $V(y)$, then the following estimates for the expected value and variance of y are obtained:

$$E(y) \doteq y(\mathbf{x}_0) + \frac{1}{2}\sum_{j=1}^{nX} [\partial^2 f(\mathbf{x}_0)/\partial x_j^2] V(x_j) \tag{2.19}$$

and

$$V(y) \doteq \sum_{j=1}^{nX} [\partial f(\mathbf{x}_0)/\partial x_j]^2 V(x_j) + \sum_{j=1}^{nX} [\partial f(\mathbf{x}_0)/\partial x_j][\partial^2 f(\mathbf{x}_0)/\partial x_j^2]\mu_3(x_j), \tag{2.20}$$

where $\mu_3(x_j)$ denotes the third central moment of x_j. As higher-order terms and correlations between the elements of $\mathbf{x}$ are included, the approximations to the expected value and variance for y rapidly become very complicated [61–64].

Differential analysis has long played a prominent role in the propagation and analysis of uncertainty [65–70]. Usually, the most difficult part of a differential analysis is determining the necessary partial derivatives. As a result, much of the research related to differential analysis has been devoted to the development of techniques for the determination of these derivatives, including adjoint techniques [71–74], Green's function techniques [75–78], and various numerical techniques [79,80]. Automatic differentiation techniques are maturing and can now be applied to quite complex programs, which greatly facilitates the implementation of derivative-based analyses [81–88].

2.3. Response surface methodology

Response surface methodology (RSM) is similar to Monte Carlo analysis except that an experimental design is used to select model input. A variety of possible designs exist, including factorial, fractional factorial, central composite, Plackett-Burman, and many more. Usually, the design selected depends on many factors, including properties of the model and the type of results desired from subsequent uncertainty and sensitivity analyses.

The experimental design results in a selection of points

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{i,nX}], \quad i = 1, 2, \ldots, nS, \tag{2.21}$$

from the $\mathscr{S}_{su}$. However, the distributions in Eq. (1.2), and hence the probability space $(\mathscr{S}_{su}, \mathfrak{S}_{su}, p_{su})$, do not play a direct role in the selection of the $\mathbf{x}_i$. Rather, these points are typically selected on the basis of the ranges of the individual x_j contained in $\mathbf{x}$ (e.g. a low, central and high value for each x_j). As a result, there is not a probabilistic weight that can be associated with each design point in Eq. (2.21) as there is with the sample elements in Eq. (2.1).

After the design points in Eq. (2.21) are selected, evaluation of f for these points creates a mapping between model input and model results of the form shown in Eq. (2.3). However, because probabilistic weights cannot be assigned to the design points, uncertainty results of the form indicated in Eqs. (2.4)–(2.6) cannot be obtained directly from these evaluations. Rather, as an intermediate step, a response surface of the form indicated in Eq. (2.8) is constructed; more complex constructions are also possible. Once constructed, this response surface can be used in a Monte Carlo simulation with the distributions in Eq. (1.2) to estimate the uncertainty in y. Or, as an alternative, expected values and variances can be determined with propagation procedures similar to those shown in Eqs. (2.10)–(2.12).

The response surface in Eq. (2.8) is analogous to the Taylor series in Eq. (2.9). Specifically, each is a linear approximation to the model $y = f(\mathbf{x})$ indicated in Eq. (1.3). As a result, sensitivity analysis in RSM can be carried out in the same manner as sensitivity analysis in differential analysis. Specifically, sensitivity measures of the type indicated in Eqs. (2.13)–(2.17) can be calculated for a regression model of the form indicated in Eq. (2.8) and derived from the mapping

$$[\mathbf{x}_i, y(\mathbf{x}_i)], \quad i = 1, 2, \ldots, nS, \tag{2.22}$$

associated with the design points in Eq. (2.21). In the context of the regression model in Eq. (2.8), the normalized

coefficients in Eqs. (2.15) and (2.17) are known as standardized regression coefficients (SRCs).

An extensive literature exists on experimental designs for use in RSM [89–101], and many examples of the use of RSM in uncertainty and sensitivity analysis exist [102–108]. In a related but somewhat different problem, RSM is widely used in optimization problems, with this area of application actually being the source from which RSM developed [109–113]. In addition, several books related to RSM are also available [114–117].

2.4. Fourier amplitude sensitivity test and Sobol' variance decomposition

The variance $V(y)$ associated with the model $y = f(\mathbf{x})$ in Eq. (1.3) is formally defined by the integral in Eq. (1.7). Although different in computational details, analyses based on both FAST [118–120] and the Sobol' variance decomposition [121] involve a decomposition of $V(y)$ into components due to individual variables and interactions between individual variables. Specifically, $V(y)$ can be decomposed into the form

$$V(y) = \sum_{1\le j\le nX} V_j + \sum_{1\le j<k\le nX} V_{jk} + \cdots + V_{1,2,\ldots,nX} \tag{2.23}$$

under the assumption that the x_j are independent, where V_j is the part of $V(y)$ due solely to x_j, V_{jk} is the part of $V(y)$ due to the interaction of x_j and x_k, V_{jkl} is the part of $V(y)$ due to the interaction of x_j, x_k and x_l, and so on up to $V_{12,\ldots,nX}$, which is the part of $V(y)$ due to the interaction of $x_1, x_2, \ldots, x_{nX}$.

Once the decomposition in Eq. (2.23) is available, various sensitivity measures such as

$$s_j = V_j/V(y) \tag{2.24}$$

$$s_{jk} = V_{jk}/V(y) \tag{2.25}$$

$$s_{jT} = \left(V_j + \sum_{\substack{l\le k<l\le nX \\ k \text{ or } l=j}} V_{kl} + \cdots + V_{12,\ldots,nX} \right) \Big/ V(y) \tag{2.26}$$

can be defined, where s_j is the fraction of $V(y)$ due solely to x_j, s_{jk} is the fraction of $V(y)$ due to the interaction of x_j and x_k, and s_{jT} is the fraction of $V(y)$ due to x_j or the interaction of x_j with other variables.

In the FAST approach, the multidimensional integrals in Eqs. (1.7) and (1.8) that define $V(y)$ and $E(y)$ are converted to the one-dimensional integrals through the construction of an appropriate space-filling curve

$$\mathbf{c}(s) = [G_1(\sin\,\omega_1 s), G_2(\sin\,\omega_2 s), \ldots, G_{nX}(\sin\,\omega_{nX} s)] \tag{2.27}$$

in $\mathscr{S}_{su}$, where the G_j and ω_j are suitably defined functions and integers, respectively. Then,

$$E(y) \doteq \frac{1}{2\pi} \int_{-\pi}^{\pi} f[\mathbf{c}(s)]\mathrm{d}s \tag{2.28}$$

and

$$V(y) \doteq \frac{1}{2\pi} \int_{-\pi}^{\pi} f^2[\mathbf{c}(s)]\mathrm{d}s - E^2(y). \tag{2.29}$$

In general, some type of numerical procedure (e.g. Monte Carlo) is required to evaluate the integrals in Eqs. (2.28) and (2.29).

The following relationship can be established by using properties of the Fourier series representation for f :

$$V(y) \doteq \sum_{k=1}^{\infty} (A_k^2 + B_k^2), \tag{2.30}$$

where

$$A_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f[\mathbf{c}(s)]\cos(ks)\mathrm{d}s$$

$$B_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f[\mathbf{c}(s)]\sin(ks)\mathrm{d}s.$$

Further, V_j can be approximated by

$$V_j \doteq \sum_{k=1}^{\infty} (A_{k\omega_j}^2 + B_{k\omega_j}^2), \tag{2.31}$$

where ω_j is the integer associated with G_j in Eq. (2.27) in the conversions from multidimensional integrals to one-dimensional integrals in Eqs. (2.28) and (2.29). Thus, the approximation

$$s_j = V_j/V(y) \doteq \sum_{k=1}^{\infty} (A_{k\omega_j}^2 + B_{k\omega_j}^2) / \sum_{k=1}^{\infty} (A_k^2 + B_k^2) \tag{2.32}$$

follows from Eqs. (2.30) and (2.31).

In analyses based on the Sobol' variance decomposition, $E(y)$ and $V(y)$ are typically approximated by Monte Carlo techniques as indicated in Eqs. (2.5) and (2.6). Further, the individual terms, V_j, V_{jk}, $V_{jkl}, \ldots, V_{12,\ldots,nX}$ in the decomposition of $V(y)$ in Eq. (2.23) are defined by multiple integrals involving the elements x_j of $\mathbf{x}$. For example,

$$V_j = \int_{\mathscr{S}_j} \left[\int_{\prod_{l\in I(-j)} \mathscr{S}_l} f(\mathbf{x}) \prod_{l\in I(-j)} d_l(x_l)\mathrm{d}x_l \right]^2 \times d_j(x_j)\mathrm{d}x_j - E^2(y) \tag{2.33}$$

$$V_{jk} = \int_{\mathscr{S}_j} \int_{\mathscr{S}_k} \left[\int_{\prod_{l\in I(-j,-k)} \mathscr{S}_l} f(\mathbf{x}) \prod_{l\in I(-j,-k)} d_l(x_l)\mathrm{d}x_l \right]^2 d_k(x_k) \times d_j(x_j)\mathrm{d}x_k\,\mathrm{d}x_j - V_j - V_k - E^2(y), \tag{2.34}$$

where, as indicated earlier, the x_j are assumed to be independent, $(\mathscr{S}_j, \mathfrak{S}_j, p_j)$ is the probability space characterizing the uncertainty in x_j, d_j is the density function associated with $(\mathscr{S}_j, \mathfrak{S}_j, p_j)$, $I(-j)$ and $I(-j,-k)$ denote the subsets of $I = \{1,2,\ldots,nX\}$ that result from the deletion of $\{j\}$ and $\{j, k\}$, respectively, and the use of the product symbol

(i.e. $\prod$) in conjunction with sets implies the concatenation of the elements of these sets. The probability space ($\mathscr{S}_{su}$, $\mathfrak{S}_{su}$, p_{su}) associated with $\mathbf{x}$ is related to the probability spaces ($\mathscr{S}_j$, $\mathfrak{S}_j$, p_j) associated with the x_j by $\mathscr{S}_{su} = \prod_{j\in I} \mathscr{S}_j$, $\mathfrak{S}_{su}$ developed from the $\mathfrak{S}_j$, $p_{su} = \prod_{j\in I} p_j$ and $d_{su} = \prod_{j\in I} d_j$.

The integrals in Eqs. (2.33) and (2.34) are quite complex and in practice must be evaluated with some type of numerical procedure (e.g. Monte Carlo [121], the Winding Stairs sampling scheme [122–124], or simplifying approximations to f [125]). Once the necessary integrals, and hence V_j and V_{jk}, are evaluated, s_j and s_{jk} can be determined as indicated in Eqs. (2.24) and (2.25).

The determination of $V_{jkl}, V_{jklm}, \ldots, V_{12,\ldots,nX}$ with either the FAST approach or Sobol' indices is very demanding computationally and typically is not done. However, relatively efficient procedures exist to evaluate the total effect sensitivity measure with both the FAST approach [126] and the Sobol' variance decomposition [127].

Additional information on the FAST approach and Sobol' variance decomposition is available in a number of publications [118–121,125–133]. Further, a conceptually equivalent approach based on analysis of variance has also been developed [134–137] and has the desirable feature of allowing correlations between the elements of $\mathbf{x}$ [135–137].

2.5. Fast probability integration

Fast probability integration (FPI) is based on the use of analytic procedures to evaluate distribution functions [138, 139]. Specifically, the following approximation procedure is used:

$$\begin{aligned}\text{prob}(y > Y) &= \int_{\mathscr{S}_{su}} \delta_Y[f(\mathbf{x})] d_{su}(\mathbf{x}) \mathrm{d}V_{su} \\ &= \int_{\mathscr{S}_{su,n}} \delta_Y[f_n(\mathbf{u})] d_{su,n}(\mathbf{u}) \mathrm{d}V_{su,n} \\ &\doteq \text{erfc}(\beta/\sqrt{2})/2, \end{aligned} \tag{2.35}$$

where (i) ($\mathscr{S}_{su,n}$, $\mathfrak{S}_{su,n}$, $p_{su,n}$) represents the probability space that results when the elements x_j of the vectors $\mathbf{x}$ associated with ($\mathscr{S}_{su}$, $\mathfrak{S}_{su}$, p_{su}) are transformed to elements u_j of $\mathbf{u} \in \mathscr{S}_{su,n}$ that are mutually independent, standardized to mean zero and standard deviation one, and normally distributed, (ii) f_n denotes the reformulation of f that uses $\mathbf{u}$ rather than $\mathbf{x}$ as its argument, and (iii) β is related to the most probable point (MPP) for which $f_n(\mathbf{u}) = Y$ as described below and erfc is the complementary error function (i.e. $\text{erfc}(x) = (2/\sqrt{\pi}) \int_x^\infty \exp(-t^2)\mathrm{d}t$).

The equality

$$f_n(\mathbf{u}) = Y \tag{2.36}$$

defines a surface in $\mathscr{S}_{su,n}$ (see Fig. 1, Ref. [138]). The MPP $\mathbf{u}_0 = [u_{10}, u_{20}, \ldots, u_{nX,0}]$ is the point on this surface that is closest to the origin in $\mathscr{S}_{su,n}$. In turn, β is given by

$$\beta = \|\mathbf{u}_0\| = [u_{10}^2 + u_{20}^2 + \cdots + u_{nX,0}^2]^{1/2} \tag{2.37}$$

and equals the distance from $\mathbf{u}_0$ to the origin. The outcome of this approach is that prob($y > Y$) is being approximated by the probability of the part of $\mathscr{S}_{su,n}$ that is cut off by a hyperplane that passes through the MPP $\mathbf{u}_0$ and is tangent to the surface defined by Eq. (2.36).

There are two major components to the implementation of an analysis based on FPI. First, the distributions indicated in Eq. (1.2) must be transformed to independent, normal distributions with mean zero and standard deviation one. Second, the MPP $\mathbf{u}_0$ associated with each probability prob($y > Y$) under consideration must be determined. This determination is typically based on search procedures using the partial derivatives of f_n with respect to the individual elements of $\mathbf{u}$. In addition, more sophisticated approximations to the surface at the MPP $\mathbf{u}_0$ than a hyperplane, and hence more sophisticated approximations to prob($y > Y$), can be developed. Although FPI is primarily used for uncertainty analysis, it can support some types of sensitivity analysis (Section 2.4, Ref. [138]).

Additional information on FPI and related techniques is available from a number of sources [138–147].

2.6. Comparison of techniques

All techniques have positive and negative features, and no single technique is optimum for all situations. In the following, the positive and negative features of the individual techniques are briefly reviewed.

Monte Carlo techniques are based on the use of a probabilistic procedure to select model input and result in a mapping between analysis inputs and analysis outcomes that is then used to produce uncertainty and sensitivity analysis results. Desirable features of Monte Carlo analysis include (i) extensive sampling from the ranges of the uncertain variables, (ii) uncertainty results that are obtained without the use of surrogate models (e.g. Taylor series in differential analysis and response surfaces in RSM), (iii) extensive modifications of, or manipulations with, the original model are not required (i.e. as is the case for the other techniques), (iv) the extensive sampling from the individual variables facilitates the identification of non-linearities, thresholds and discontinuities, (v) a variety of sensitivity analysis procedures are available, and (vi) the approach is conceptually simple, widely used, and easy to explain. The major drawback is computational cost. This is especially the case if long-running models are under consideration or probabilities very close to zero or one must be estimated.

Differential analysis is based on developing a Taylor series approximation to the model under consideration. Desirable properties of differential analysis include (i) the effects of small perturbations away from the base-case value at which the Taylor series is developed are revealed, (ii) uncertainty and sensitivity analyses based on variance propagation are straightforward once the Taylor series is developed, (iii) techniques (e.g. adjoint, Green's function,

specialized compilers) exist to facilitate the calculation of derivatives, and (iv) the approach has been widely studied and applied. There are two primary drawbacks: (i) differential analysis is inherently local, and (ii) a differential analysis can be difficult to implement and can require large amounts of human and/or computational time.

Response surface methodology (RSM) is based on using an experimental design to select model input and then developing a response surface replacement for the original model that is used in subsequent uncertainty and sensitivity analyses. Desirable properties of RSM include (i) complete control over the structure of the model input through the experimental design selected for use, (ii) near optimum choice for a model whose predictions are known to be a linear or quadratic function of the input variables, (iii) uncertainty and sensitivity analyses are straightforward once the necessary response surface replacement has been developed, and (iv) experimental designs for use in RSM have been widely studied. Drawbacks to RSM include (i) difficulty of developing an appropriate experimental design, (ii) use of a limited number of values for each input variable, (iii) possible need for a large number of design points, (iv) difficulties in detecting thresholds, discontinuities and nonlinearities, (v) difficulty in including correlations and restrictions between input variables, and (vi) difficulty in constructing an appropriate response surface approximation to the model under consideration.

The FAST approach and Sobol' variance decomposition are based on a direct decomposition of variance into the parts contributed by individual variables. Desirable properties of the FAST approach and Sobol' variance decomposition include (i) full range of each input variable is explored, (ii) estimation of expected value and variance is by direct calculation rather than by use of a surrogate model, (iii) fractional contribution of each variable to total variance is determined, (iv) effects of variable interactions can be determined, (v) sensitivity analysis is not predicated on a search for linear or monotonic relationships, and (vi) modifications to the original model are not required. Drawbacks include (i) the mathematics is complicated and difficult to explain, (ii) the approaches are not widely known and applied, (iii) evaluating the required integrals can be both complex and computationally demanding, and (iv) correlations cannot be imposed on the input variables.

Fast probability integration (FPI) is based on the use of analytic procedures to evaluate distribution functions. The desirable feature of FPI is that it allows the estimation of the tails of a distribution without the estimation of the full distribution. This has the potential to require less computation than the use of Monte Carlo procedures to estimate the same tail probabilities. Less desirable features are that (i) the underlying mathematics is complicated and difficult to explain, (ii) the calculation of the partial derivatives required in the approach can be computationally demanding, and (iii) the approach is not appropriate for the calculation of full distributions or the consideration of distributions for a large number of different variables. Further, the approach is primarily one of uncertainty analysis and lacks associated sensitivity analysis procedures.

This review considers the use of Monte Carlo techniques in general and Latin hypercube sampling in particular in analyses that involve the propagation of uncertainty through complex systems. Although a variety of techniques exist for the propagation of uncertainty as previously indicated, Monte Carlo techniques provide the most effective approach to the propagation and analysis of uncertainty in many situations for various combinations of the following reasons: (i) large uncertainties are often present and a sampling-based approach provides a full coverage of the range of each uncertain variable, (ii) modification of the model is not required, (iii) direct estimates of distribution functions are provided, (iv) analyses are conceptually simple and logistically easy to implement, (v) analysis procedures can be developed that allow the propagation of results through systems of linked models, and (vi) a variety of sensitivity analysis procedures are available. Latin hypercube sampling is often the preferred sampling procedure in Monte Carlo analyses due to the efficient manner in which it stratifies across the range of each sampled variable.

3. Random, stratified and Latin hypercube sampling

3.1. Description of sampling techniques

In Monte Carlo analysis, some type of sampling procedure must be used to generate the sample in Eq. (2.1). The simplest procedure is random sampling. With random sampling from uncorrelated variables, each sample element is generated independently of all other sample elements, and the probability that this element will come from a particular subset of $\mathscr{S}_{su}$ (i.e. $\mathscr{E} \in \mathfrak{S}_{su}$) is equal to the probability of that subset (i.e. $p_{su}(\mathscr{E})$).

The nature of a random sample will be illustrated for $\mathbf{x} = [U, V]$, U assigned a uniform distribution on [0, 10], V assigned a triangular distribution and a mode of 8 on [0, 10], and $nS = 5$. The sample is generated by independently sampling five random numbers $RU(1), RU(2), \ldots, RU(5)$ from a uniform distribution on [0, 1] and then using the CDF for U to obtain five values $U(1), U(2), \ldots, U(5)$ for U (Fig. 3a). Similarly, random sampling is again used to obtain an additional five independent random numbers $RV(1), RV(2), \ldots, RV(5)$ from a uniform distribution on [0, 1], and the CDF for V is used to obtain five values $V(1), V(2), \ldots, V(5)$ for V (Fig. 3b). Then,

$$\mathbf{x}_i = [U(i), V(i)], \quad i = 1, 2, \ldots, 5, \tag{3.1}$$

constitutes a random sample of size $nS = 5$ generated in consistency with the distributions assigned to U and V

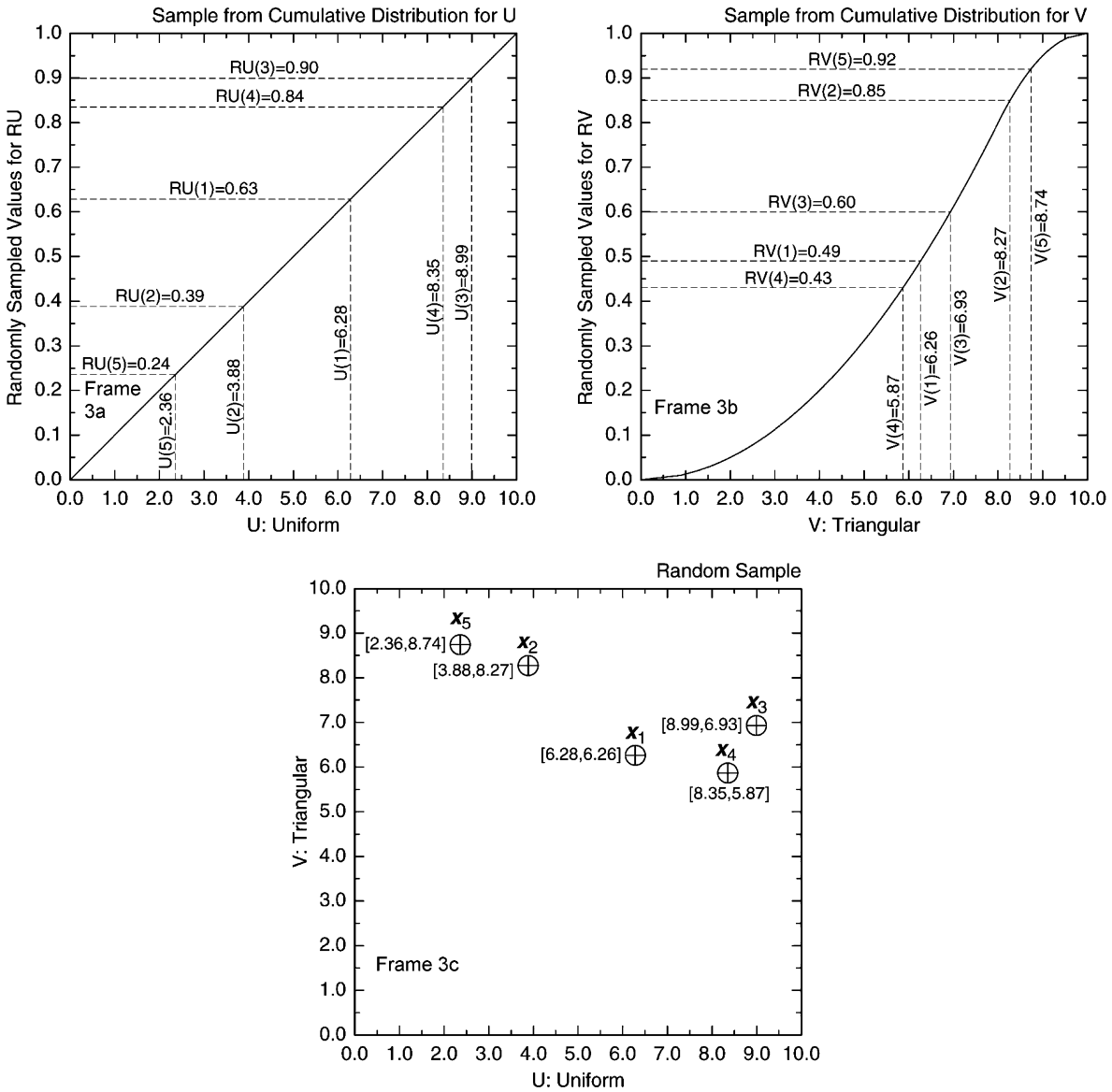

Fig. 3. Generation of a random sample of size $nS = 5$ from $\mathbf{x} = [U, V]$ with U uniform on [0, 10] and V triangular on [0, 10] with a mode of 8.

(Fig. 3c). The generation of a random sample

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{i,nX}], \quad i = 1, 2, \ldots, nS, \tag{3.2}$$

when $\mathbf{x}$ has dimension $nX > 2$ is carried out in an analogous manner.

The generation of a random sample in multiple dimensions ultimately depends on being able to generate uniformly distributed random numbers from the interval [0, 1]. The generation of such random numbers has been widely studied and discussed [57,148–150]. As an aside, such numbers are often called pseudorandom numbers because they are generated by reproducible algorithmic processes rather than in a truly random manner. For this presentation, the capability to generate random numbers is taken for granted and discussed no further.

With random sampling, there is no assurance that a sample element will be generated from any particular subset of the sample space $\mathscr{S}_{\text{su}}$. In particular, important subsets of $\mathscr{S}_{\text{su}}$ with low probability but high consequences are likely to be missed. Stratified sampling, or importance sampling as it is also sometimes called, provides a way to mitigate this problem by specifying subsets of $\mathscr{S}_{\text{su}}$ from which sample elements will be selected. Specifically, $\mathscr{S}_{\text{su}}$ is exhaustively subdivided into a collection $\mathscr{E}_1, \mathscr{E}_2, \ldots, \mathscr{E}_{nI}$ of disjoint subsets (i.e. $\bigcup_{k=1}^{nI} \mathscr{E}_k = \mathscr{S}_{\text{su}}$ and $\mathscr{E}_p \cap \mathscr{E}_q = \varnothing$ for $p \neq q$) (Fig. 4). The $\mathscr{E}_k$ constitute the strata associated with the sampling procedure. Then, the corresponding sample (i.e. the stratified or importance sample)

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{i,nX}], \quad i = 1, 2, \ldots, nS = \sum_{k=1}^{nI} nI_k, \tag{3.3}$$

is obtained by randomly sampling nI_k sample elements from strata $\mathscr{E}_k$. The preceding sampling is carried out conditional

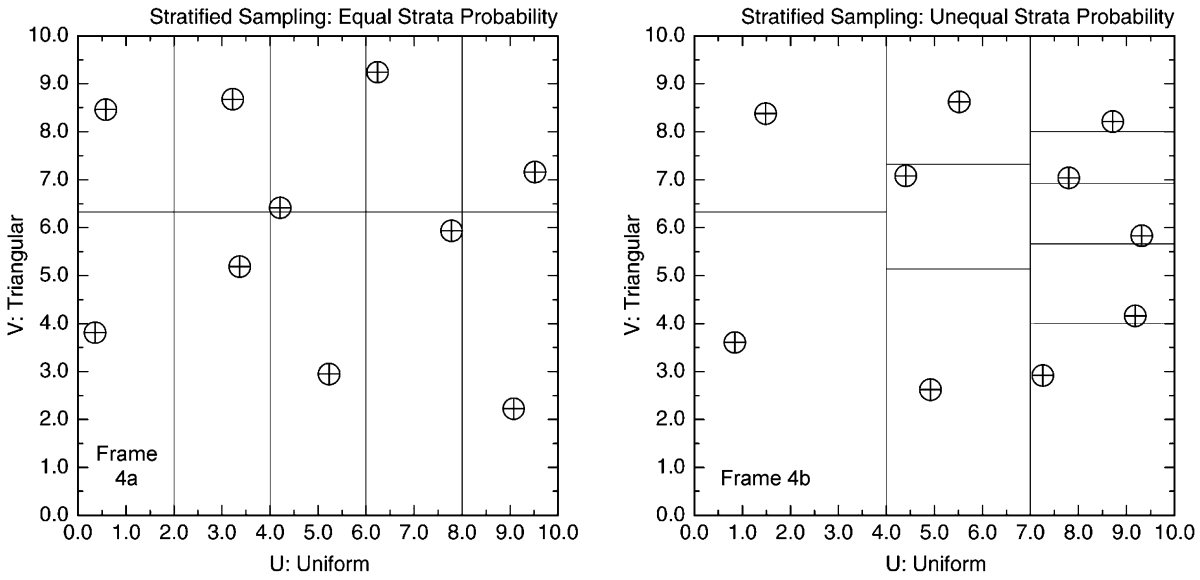

Fig. 4. Generation of a stratified sample of size $nS = 10$ with one random sample per strata (i.e. $nI_k = 1$) from $\mathbf{x} = [U, V]$ with U uniform on [0, 10] and V triangular on [0, 10] with a mode of 8: (a) equal strata probability (i.e. $p_{su}(\mathscr{E}_k) = 0.1$), and (b) unequal strata probability (i.e. $p_{su}(\mathscr{E}_k) = 0.2, 0.2, 0.1, 0.1, 0.1, 0.06, 0.06, 0.06, 0.06, 0.06$).

on the restriction of $\mathbf{x}$ to $\mathscr{E}_k$. Further, if $\mathbf{x}_i \in \mathscr{E}_k$, then the corresponding weight w_i for use in probabilistic calculations is given by $w_i = p_{su}(\mathscr{E}_k)/nI_k$. In most applications, $nI_k = 1$, and so the sample size nS is equal to the number of strata and $w_i = p_{su}(\mathscr{E}_k)$ for $\mathbf{x}_i \in \mathscr{E}_k$.

Stratified sampling has the advantage of forcing the inclusion of specified subsets of $\mathscr{S}_{su}$ while maintaining the probabilistic character of random sampling. Indeed, it can be argued that stratified sampling is always the best procedure to use when enough information is available for its appropriate implementation. A major problem associated with stratified sampling is the necessity of defining the strata $\mathscr{E}_1, \mathscr{E}_2, \ldots, \mathscr{E}_{nI}$ and also calculating their probabilities. Both of these requirements are avoided when random sampling is used. When the dimensionality of $\mathscr{S}_{su}$ is high, the determination of strata and strata probabilities becomes a major undertaking. The event tree and fault procedures that underlie many large analyses can be viewed as algorithms to determine the strata and strata probabilities for use in a stratified sampling procedure. These determinations are further complicated when many analysis outcomes are under consideration (i.e. when $\mathbf{y}$ in Eq. (1.1) is of high dimension); in particular, strata definitions that are appropriate for one analysis outcome may be inappropriate for other analysis outcomes. A compounding problem is that all the analysis outcomes that will be studied in the course of an analysis may not even be known at the beginning of the analysis.

Latin hypercube sampling can be viewed as a compromise procedure that incorporates many of the desirable features of random sampling and stratified sampling and also produces more stable analysis outcomes than random sampling. Like random and stratified sampling, Latin hypercube sampling is a probabilistic procedure in the sense that a weight (i.e. $w_i = 1/nS$) can be associated with each sample element that can be used in probabilistic calculations (i.e. in the estimation of the integrals in Eqs. (1.4)–(1.8)). Like random sampling, the implementation of Latin hypercube sampling is easier than the implementation of stratified sampling because it is not necessary to determine strata and strata probabilities. However, Latin hypercube sampling does have the property of densely stratifying across the range of each element of $\mathbf{x}$, which is a property closer to those possessed by stratified sampling. Thus, Latin hypercube sampling displays properties between random sampling, which involves no stratification, and stratified sampling, which stratifies on $\mathscr{S}_{su}$.

Latin hypercube sampling operates in the following manner to generate a sample of size nS from $\mathbf{x} = [x_1, x_2, \ldots, x_{nX}]$ in consistency with the distributions $D_1, D_2, \ldots, D_{nX}$ indicated in Eq. (1.2) (i.e. in consistency with the probability space $(\mathscr{S}_{su}, \Xi_{su}, p_{su})$). The range of each variable (i.e. the x_j) is exhaustively divided into nS disjoint intervals of equal probability and one value is selected at random from each interval. The nS values thus obtained for x_1 are paired at random without replacement with the nS values obtained for x_2. These nS pairs are combined in a random manner without replacement with the nS values of x_3 to form nS triples. This process is continued until a set of nS nX-tuples is formed. These nX-tuples are of the form

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{i,nX}], \quad i = 1, 2, \ldots, nS, \tag{3.4}$$

and constitute the Latin hypercube sample (LHS). The individual x_j must be independent for the preceding construction procedure to work; a method for generating Latin hypercube and random samples from correlated variables has been developed by Iman and Conover [151]

and will be discussed in Section 5.1. Latin hypercube sampling is an extension of quota sampling [152] and can be viewed as an n-dimensional randomized generalization of Latin square sampling (Ref. [153], pp. 206–209).

The generation of an LHS is illustrated for $\mathbf{x} = [U, V]$ and $nS = 5$ (Fig. 5). The ranges of U and V are subdivided into five intervals of equal probability, with this subdivision represented by the lines that originate at 0.2, 0.4, 0.6 and 0.8 on the ordinates of Figs. 5a and b, extend horizontally to the CDFs, and then drop vertically to the abscissas to produce the five indicated intervals. Random values $U(1), U(2), \ldots, U(5)$ and $V(1), V(2), \ldots, V(5)$ are then sampled from these intervals. The sampling of these random values is implemented by (i) sampling $RU(1)$ and $RV(1)$ from a uniform distribution on [0, 0.2], $RU(2)$ and $RV(2)$ from a uniform distribution on [0.2, 0.4], and so on, and then (ii) using the CDFs to identify (i.e. sample) the corresponding U and V values, with this identification represented by the dashed lines that originate on the ordinates of Figs. 5a and b, extend horizontally to the CDFs, and then drop vertically to the abscissas to produce $U(1), U(2), \ldots, U(5)$ and $V(1), V(2), \ldots, V(5)$. The generation of the LHS is then completed by randomly pairing (without replacement) the resulting values for U and V. As this pairing is not unique, many possible LHSs can result, with the LHS in Fig. 5c resulting from the pairings $[U(1), V(4)]$, $[U(2), V(2)]$, $[U(3), V(1)]$, $[U(4), V(5)]$, $[U(5), V(3)]$ and the LHS in Fig. 5d resulting from the pairings $[U(1), V(5)]$, $[U(2), V(1)]$, $[U(3), V(3)]$, $[U(4), V(4)]$, $[U(5), V(2)]$.

The generation of an LHS for $nS > 2$ proceeds in a manner similar to that shown in Fig. 5 for $nV = 2$. The sampling of the individual variables for $nS > 2$ takes place

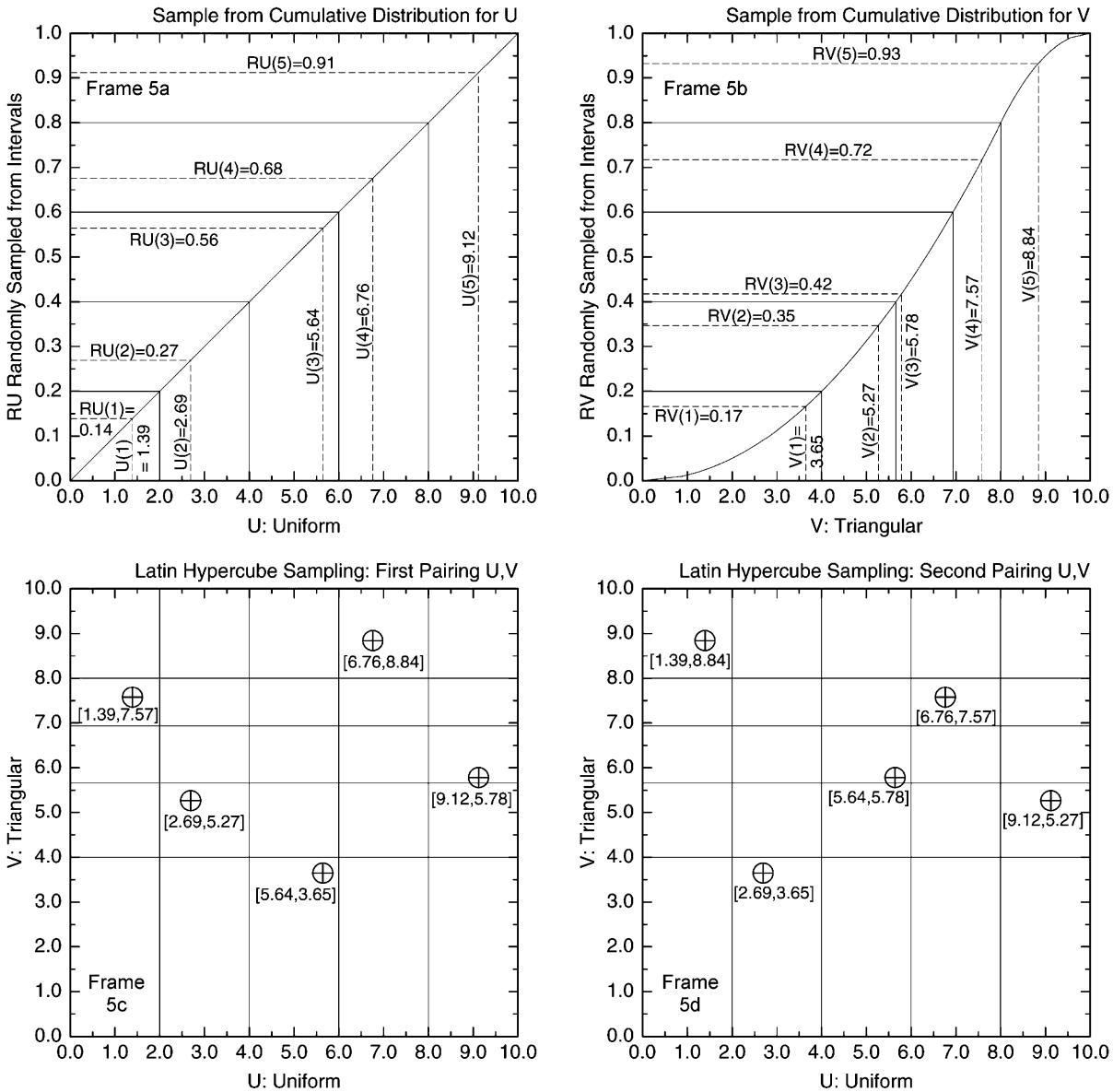

Fig. 5. Example of Latin hypercube sampling to generate a sample of size $nS = 5$ from $\mathbf{x} = [U, V]$ with U uniform on [0, 10] and V triangular on [0, 10] with a mode of 8.

in the same manner as shown in Figs. 5a and b. However, the nX variables define an nX-dimensional solid rather than a two-dimensional rectangle in the plane. Thus, Figs. 5c and d would involve a partitioning of an nX-dimensional solid rather than a rectangle.

3.2. Properties of sampling techniques

Random sampling, stratified sampling and Latin hypercube sampling are now discussed and compared. This discussion is derived from the study by McKay et al. [31]. For notational convenience, a single element y of the vector $\mathbf{y}$ in Eq. (1.1) is considered.

The following estimator is widely used in conjunction with random sampling:

$$T(y_1, y_2, \ldots, y_{nS}) = (1/nS)\sum_{i=1}^{nS} g(y_i), \tag{3.5}$$

where $y_i = f(\mathbf{x}_i)$ for the random sample appearing in Eq. (3.2) and g is an arbitrary function. If $g(y) = y$, then T represents the sample mean, which is used to estimate the expected value $E(y)$ of y. If $g(y) = y^r$, then T represents an estimate for the rth sample moment, which is used in obtaining an estimate for the corresponding population moment. If $g(y) = 1$ for $y \le Y$ and $g(y) = 0$ otherwise, then T is an estimate of the quantile on the distribution function of y associated with y. Let Υ denote the expected value for the population of Ts that results from repeated calculations with independent random samples of size nS from $\mathbf{x}$. McKay et al. [31] show that both stratified sampling and Latin hypercube sampling yield unbiased estimates for Υ, which is also the case for random sampling. That is, the expected value of repeated calculations of T with either sampling method is Υ.

For notational convenience, let T_R, T_S and T_L represent estimates of Υ (i.e. values of T calculated as shown in Eq. (3.5)) obtained with a random sample of size nS, a stratified sample of size nS with all strata of equal probability and one random selection per strata, and an LHS of size nS, respectively. Then, as shown by McKay et al. [31],

$$\mathrm{Var}(T_S) \le \mathrm{Var}(T_R), \tag{3.6}$$

where Var represents the variance of T_S and T_R under repeated estimations. No direct means of comparing the variance of T_L and T_R appears to be known. However, the following result has been established by McKay et al. [31].

Theorem 3.1. *If $y = f(x_1, x_2, \ldots, x_{nX})$ is monotonic in each of the x_j and $g(y)$ is a monotonic function of y, then*

$$\mathrm{Var}(T_L) \le \mathrm{Var}(T_R). \tag{3.7}$$

As indicated earlier, uncertainty analysis generally involves estimating the mean, variance and distribution function for the particular dependent variable under consideration. Estimates for these quantities with random sampling, stratified sampling, and Latin hypercube sampling are now considered. For each sampling method, the form for the estimator of the expected value of y is given by

$$\bar{y} = \hat{E}(y) = (1/nS)\sum_{i=1}^{nS} y_i, \tag{3.8}$$

where $y_i = f(\mathbf{x}_i)$. To obtain this representation for the stratified sample, it is assumed that $\mathbf{x}_i$ comes from stratum $\mathscr{E}_i$, $p_{su}(\mathscr{E}_i) = 1/nS$, and $nI_i = 1$. The symbols $\bar{y}_R, \bar{y}_S$ and $\bar{y}_L$ are used to represent the value obtained in Eq. (3.8) with random sampling, stratified sampling, and Latin hypercube sampling, respectively. Each of $\bar{y}_R$, $\bar{y}_S$ and $\bar{y}_L$ is an unbiased estimator of $E(y)$.

The goodness of an unbiased estimator can be measured by its variance. As shown in McKay et al. [31],

$$\mathrm{Var}(\bar{y}_R) = (1/nS)\mathrm{Var}(y), \tag{3.9}$$

$$\mathrm{Var}(\bar{y}_S) = \mathrm{Var}(\bar{y}_R) - (1/nS^2)\sum_{i=1}^{nS} (\mu_i - \mu)^2, \tag{3.10}$$

and

$$\mathrm{Var}(\bar{y}_L) = \mathrm{Var}(\bar{y}_R) + \frac{nS-1}{nS^{nX+1}(nS-1)^{nX}}\sum_{\mathscr{R}} (\mu_r - \mu)(\mu_s - \mu), \tag{3.11}$$

where

$$\mu = E(y), \tag{3.12}$$

$$\mu_i = E(y|\mathbf{x} \in \mathscr{E}_i) \tag{3.13}$$

in Eq. (3.10) for the stratified sample,

$$\mu_r = E(y|\mathbf{x} \in \mathrm{cell}\, r) \tag{3.14}$$

in Eq. (3.11) for the LHS, and $\mathscr{R}$ in Eq. (3.11) denotes the restricted space of all pairs (μ_r, μ_s) for which the associated cells have no coordinates in common. The cells being referred to in conjunction with Latin hypercube sampling in Eq. (3.11) are the nS^{nX} possible combinations of intervals of equal probability used in the construction of the sample. Each cell can be labeled by a set of coordinates

$$\mathbf{m}_r = [m_{r1}, m_{r2}, \ldots, m_{r,nX}], \tag{3.15}$$

where m_{rj} is the interval number for variable x_j associated with cell r, $r = 1,2,\ldots,nS^{nX}$. The statement that cells r and s have no coordinate in common means that $m_{rj} \ne m_{sj}$ for $j = 1,2,\ldots,nX$.

Comparison of Eqs. (3.9) and (3.10) shows that

$$\mathrm{Var}(\bar{y}_S) \le \mathrm{Var}(\bar{y}_R). \tag{3.16}$$

The relationship between $\mathrm{Var}(\bar{y}_R)$ and $\mathrm{Var}(\bar{y}_L)$ is not easily ascertained by comparing Eqs. (3.9) and (3.11). However, the previously stated theorem by McKay et al. [31] (Theorem 3.1) implies that

$$\mathrm{Var}(\bar{y}_L) \le \mathrm{Var}(\bar{y}_R) \tag{3.17}$$

when $y = f(x_1, x_2, \ldots, x_{nX})$ is monotonic in each of the x_j. In the example presented in McKay et al. [31], the sampling variability in $\bar{y}_L$ (i.e. $\mathrm{Var}(\bar{y}_L)$) was considerably less than that for $\bar{y}_R$ and $\bar{y}_S$.

For each sampling method, the form for the estimator of the variance of y is given by

$$S^2 = (1/nS)\sum_{i=1}^{nS}(y_i - \bar{y})^2 \tag{3.18}$$

and its expectation is given by

$$E(S^2) = \mathrm{Var}(y) - \mathrm{Var}(\bar{y}), \tag{3.19}$$

where $\bar{y}$ is $\bar{y}_R$, $\bar{y}_S$ or $\bar{y}_L$, depending on which sampling technique is in use. For convenience, S_R^2, S_S^2 and S_L^2 are used to represent the values obtained in Eq. (3.18) for random sampling, stratified sampling (equal probability strata), and Latin hypercube sampling.

For the random sample, $nS\ S_R^2/(nS-1)$ is an unbiased estimator of the variance of y. The bias in the case of stratified sampling is unknown. However, it follows from Eqs. (3.9), (3.16) and (3.19) that

$$[(nS-1)/nS]\mathrm{Var}(y) \le E(S_S^2) \le \mathrm{Var}(y). \tag{3.20}$$

The bias in S_L^2 is also unknown. However, in a derivation analogous to the one used for Eq. (3.20), it follows from Eqs. (3.9), (3.17) and (3.19) that

$$[(nS-1)/nS]\mathrm{Var}(y) \le E(S_L^2) \le \mathrm{Var}(y) \tag{3.21}$$

when $y = f(x_1, x_2, \ldots, x_{nX})$ is monotonic in each of the x_j. In the example given in McKay et al. [31], S_L^2 was found to have little bias and considerably less sampling variability than either random or stratified sampling.

For each sampling method, the form for the estimator of the distribution function of y is given by

$$G(y) = (1/nS)\sum_{i=1}^{nS} u(y - y_i) \tag{3.22}$$

where $u(z) = 1$ if $z \ge 0$ and $u(z) = 0$ otherwise. More specifically, $G(y)$ is the estimator for the quantile on the distribution function associated with y. The locus of points $(y, G(y))$ is the empirical distribution function associated with $y_1, y_2, \ldots, y_{nS}$. Since Eq. (3.22) is of the form shown in Eq. (3.5), the expected value of $G(y)$ is the same under all three sampling plans. Under random sampling, $G(y)$ is an unbiased estimator for the distribution function of y, and so stratified and Latin hypercube sampling also provide unbiased estimates.

As shown in McKay et al. [31], the variances for the estimators in Eq. (3.22) are given by

$$\mathrm{Var}[G_R(y)] = (1/nS)D(y)[1 - D(y)], \tag{3.23}$$

$$\mathrm{Var}[G_S(y)] = \mathrm{Var}[G_R(y)] - (1/nS^2)\sum_{i=1}^{nS}[D_i(y) - D(y)]^2, \tag{3.24}$$

and

$$\mathrm{Var}[G_L(y)] = \mathrm{Var}[G_R(y)] + \frac{nS-1}{nS^{nX+1}(nS-1)^{nX}} \times \sum_{\mathscr{R}}[D_r(y) - D(y)][D_s(y) - D(y)], \tag{3.25}$$

where G_R, G_S and G_L represent the estimator in Eq. (3.22) with random, stratified and Latin hypercube sampling, respectively, D represents the true distribution function for y, D_i and D_r represent the distribution function for y conditional on $\mathbf{x}$ belonging to stratum i or cell r as appropriate (see Eqs. (3.13) and (3.14)), and $\mathscr{R}$ represents the same restricted space that it did in Eq. (3.11).

The equality in Eq. (3.24) implies that

$$\mathrm{Var}[G_S(y)] \le \mathrm{Var}[G_R(y)]. \tag{3.26}$$

Thus, the variance in estimating $D(y)$ with stratified sampling is less than that with random sampling. The relationship between $\mathrm{Var}[G_L(y)]$ and $\mathrm{Var}[G_R(y)]$ is not readily seen by comparing Eqs. (3.23) and (3.25). In the example given in McKay et al. [31], the sampling variability in $G_L(y)$ (i.e. $\mathrm{Var}[G_L(y)]$) was found to be considerably less than that in $G_R(y)$ and $G_S(y)$.

The comparisons involving random sampling, stratified sampling and Latin hypercube sampling discussed so far have all been for samples of a fixed size nS. Stein [154] has derived asymptotic comparisons of the variability of estimates T_R and T_L of T obtained with random sampling and Latin hypercube sampling, respectively, under the assumption that the x_js are independent. In particular, Stein found that the inequality

$$\mathrm{Var}[T_L(y_1, y_2, \ldots, y_{nS})] < \mathrm{Var}[T_R(y_1, y_2, \ldots, y_{nS})] \tag{3.27}$$

can be expected to hold for sufficiently large sample sizes nS for most models.

A more explicit statement of Stein's result requires some additional notation. Let $\mathscr{S}_{su,j}$, $j = 1, 2, \ldots, nX$, represent the sample space for x_j, and let $d_{su,j}$ represent the corresponding density function, with both $\mathscr{S}_{su,j}$ and $d_{su,j}$ deriving from the distribution D_j indicated in Eq. (1.2). Further, let $I = \{1, 2, \ldots, nX\}$, $I(-j) = I - \{j\}$, $d_{su}(\mathbf{x}) = \prod_{j\in I} d_{su,j}(x_j)$, and $d_{su,-j}(\mathbf{x}) = \prod_{k\in I(-j)} d_{su,k}(x_k)$. The representation of $d_{su}(\mathbf{x})$ and $d_{su,-j}(\mathbf{x})$ as products involving $d_{su,j}(x_j)$ is possible because the x_js are assumed to be independent.

Stein's result is based on the following decomposition of $g[f(\tilde{\mathbf{x}})]$:

$$g[f(\tilde{\mathbf{x}})] = \mu + \sum_{j=1}^{nX}\alpha_j(\tilde{\mathbf{x}}) + r(\tilde{\mathbf{x}}), \tag{3.28}$$

where

$\tilde{\mathbf{x}} = [\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_{nX}]$ is an arbitrary element of $\mathscr{S}_{su}$,

$\mu = \int_{\mathscr{S}_{su}} g[f(\mathbf{x})]d_{su}(\mathbf{x})\mathrm{d}V_{su}$,

$$\mathscr{S}_{\mathrm{su},-j}(x) = \{\mathbf{x} | \mathbf{x} \in \mathscr{S}_{\mathrm{su}} \text{ and } x_j = x\},$$

$$\alpha_j(\tilde{\mathbf{x}}) = \int_{\mathscr{S}_{\mathrm{su},-j}(\tilde{x}_j)} \{g[f(\mathbf{x})] - \mu\} d_{\mathrm{su},-j}(\mathbf{x}) \mathrm{d}V_{\mathrm{su},-j},$$

$\mathrm{d}V_{\mathrm{su},-j}$ represents an increment of volume from $\mathscr{S}_{\mathrm{su},-j}(\tilde{x}_j)$, and $r(\tilde{\mathbf{x}})$ is formally defined by

$$r(\tilde{\mathbf{x}}) = g[f(\tilde{\mathbf{x}})] - \mu - \sum_{j=1}^{nX} \alpha_j(\tilde{\mathbf{x}}). \tag{3.29}$$

The function $\alpha_j(\tilde{\mathbf{x}})$ characterizes the 'main effect' of the element $\tilde{x}_j$ of $\tilde{\mathbf{x}}$, and the function $r(\tilde{\mathbf{x}})$ characterizes the nonadditive component of $g[f(\tilde{\mathbf{x}})]$. As an aside, this decomposition also underlies the procedures introduced in Section 2.4. The following result is proved by Stein (Ref. [154], Corollary 1, p. 145).

Theorem 3.2. *If* $\int_{\mathscr{S}_{\mathrm{su}}} g^2[f(\mathbf{x})] d_{\mathrm{su}}(\mathbf{x}) \mathrm{d}V_{\mathrm{su}}$ *is finite, then*

$$\mathrm{Var}[T_{\mathrm{L}}(y_1, y_2, \ldots, y_{nS})] = \int_{\mathscr{S}_{\mathrm{su}}} r^2(\mathbf{x}) d_{\mathrm{su}}(\mathbf{x}) \mathrm{d}V_{\mathrm{su}}/nS + o(nS^{-1}), \tag{3.30}$$

where the notation $F(nS^{-1}) = o(nS^{-1})$ *indicates that* $F(nS^{-1})/nS^{-1} \to 0$ *as* $nS \to \infty$ (*Ref.* [30], *p. xv*).

The corresponding variance associated with random sampling is given by

$$\begin{aligned}\mathrm{Var}[T_{\mathrm{R}}(y_1, y_2, \ldots, y_{nS})] &= \int_{\mathscr{S}_{\mathrm{su}}} \{g[f(\mathbf{x})] - \mu\}^2 d_{\mathrm{su}}(\mathbf{x}) \mathrm{d}V_{\mathrm{su}}/nS \\ &= \int_{\mathscr{S}_{\mathrm{su}}} r^2(\mathbf{x}) d_{\mathrm{su}}(\mathbf{x}) \mathrm{d}V_{\mathrm{su}}/nS \\ &\quad + \sum_{j=1}^{nX} \alpha_j^2(\mathbf{x}) d_{\mathrm{su}}(\mathbf{x}) \mathrm{d}V_{\mathrm{su}}/nS, \end{aligned} \tag{3.31}$$

with the second equality following from Eq. (3.28) and the equalities

$$0 = \int_{\mathscr{S}_{\mathrm{su},j}} \alpha_j(\mathbf{x}) d_{\mathrm{su},j}(x_j) \mathrm{d}x_j \tag{3.32}$$

for $j = 1, 2, \ldots, nX$ and

$$0 = \int_{\mathscr{S}_{\mathrm{su},-j}(x_j)} r(\mathbf{x}) d_{\mathrm{su},-j}(\mathbf{x}) \mathrm{d}V_{\mathrm{su},-j} \tag{3.33}$$

for $x_j \in \mathscr{S}_{\mathrm{su},j}$ and $j = 1, 2, \ldots, nX$. Thus, above some sample size, Latin hypercube sampling results in estimates for T with lower variance than random sampling unless all the main effects $\alpha_j(\mathbf{x})$, $j = 1, 2, \ldots, nX$, are zero (Theorem 3.2).

For sufficiently large sample sizes, $T_{\mathrm{L}} - Y$ has a distribution that is approximately normal, where Y is the expected value of T_{L}. Specifically, the following result has been established by Owen [155].

Theorem 3.3. *If* $g[f(\mathbf{x})]$ *is bounded, then* $nS^{1/2}(T_{\mathrm{L}} - Y)$ *converges in distribution to a normal distribution with mean zero and variance*

$$\int_{\mathscr{S}_{\mathrm{su}}} r^2(\mathbf{x}) d_{\mathrm{su}}(\mathbf{x}) \mathrm{d}V_{\mathrm{su}}$$

as nS increases (*see Ref.* [156], *Section 1.4, for formal definition of convergence in distribution*).

In practice, most models satisfy the boundedness condition imposed on $g[f(\mathbf{x})]$. Thus, in concept, the preceding result can be used to place confidence intervals on results obtained with Latin hypercube sampling. In practice, determining how large nS must be for approximate normality to hold can be difficult (Theorem 3.3).

Additional results on variance reduction associated with Latin hypercube sampling and further references are given in several recent papers [157,158]. Also, a number of references related to the theoretical development of Latin hypercube sampling are given at the end of Section 5.1.

3.3. Historical development of Latin hypercube sampling

The introduction of Latin hypercube sampling can be traced to concerns in the reactor safety community over the treatment of uncertainty in analyses related to the safety of nuclear power plants. In particular, the Reactor Safety Study [159] was published by US Nuclear Regulatory Commission (NRC) in 1975 and widely praised for its advancement of the state of probabilistic risk assessment (PRA) [160]. However, it was also criticized for inadequately representing the uncertainty in its results [160]. This led to an active interest on the part of the NRC and its contractors in the propagation of uncertainty through models for complex systems.

In this environment, Latin hypercube sampling was conceived of by W.J. Conover (the original, unpublished manuscript documenting this work is reproduced in App. A in Ref. [161]) and formally published in conjunction with colleagues at Los Alamos Scientific Laboratory [31]. The first applications of Latin hypercube sampling were in the analysis of loss of coolant accidents (LOCAs) in the context of reactor safety [162,163]. R.L. Iman, a student of Conover's and a staff member at Sandia National Laboratories, recognized the potential of Latin hypercube sampling and became an early and active proponent of its use. Among his contributions was to write the first widely distributed program for Latin hypercube sampling [164,165]. A brief description of the early development of Latin hypercube sampling was prepared by Iman in 1980 (this unpublished description is reproduced in App. B in Ref. [161]).

Much of the early use of Latin hypercube sampling was in programs related to radioactive waste disposal carried out at Sandia National Laboratories for the NRC [166–168]. In addition, the NRC also supported work on Latin hypercube sampling and associated sensitivity analysis techniques as part of its MELCOR project to develop a new suite of models for use in performing reactor safety studies [169–171].

In the mid 1980s, the NRC decided to reassess the results obtained in the Reactor Safety Study, with particular

attention to be paid to the assessment and propagation of uncertainty. This study, often referred to as NUREG-1150 after its report number, was a very large analysis and probably the largest integrated analysis of any system carried out in the 1980s [172,173]. As part of the NUREG-1150 analyses, Latin hypercube sampling was used in the propagation of uncertainty through PRAs for 5 nuclear power plants [174–178]. In addition to the extensive technical report literature documenting these PRAs, summaries are also available in the journal literature [173, 179–183]. Subsequent to NUREG-1150, Latin hypercube sampling was used in a very extensive PRA for the LaSalle nuclear power station [184–187].

After the NUREG-1150 analyses, the next large project to make use of Latin hypercube sampling involved performance assessment (PA) for the Waste Isolation Pilot Plant (WIPP), which was under development by the US Department of Energy (DOE) for the geologic disposal of transuranic radioactive waste [188,189]. Latin hypercube sampling was used in several PAs for the WIPP, including the PA that supported the DOE's successful compliance certification application (CCA) to the US Environmental Protection Agency (EPA) for the WIPP [190,191]. With its certification, the WIPP became the first operational facility in the United States for the geologic disposal of radioactive waste. As an aside, EPA staff members charged with writing regulations for the geologic disposal of radioactive waste were acquainted with, and influenced by, uncertainty analyses performed with Latin hypercube sampling, with the result that the final regulations developed for the WIPP mandated an uncertainty propagation of the type for which Latin hypercube sampling is well suited [192–195].

At present, the largest project that is making use of Latin hypercube sampling is the Yucca Mountain Project (YMP) to develop a deep geologic disposal facility for high level radioactive waste at Yucca Mountain, Nevada [196–198]. This project is both large and controversial. It is also a very important project that has been much in the news recently and is likely to get even more attention in the near future for various reasons. Another large project that is currently using Latin hypercube sampling is the System Assessment Capability (SAC) program for the Hanford Site [199,200].

The preceding background discussion has concentrated on the large analyses that have used Latin hypercube sampling. However, Latin hypercube sampling has also been used in smaller analyses in a variety of fields (e.g. Refs. [33–39,41–44,201–215]). A recent check (Sept. 9, 2001) of SciSearch shows 330 citations to the original article on Latin hypercube sampling [31], with the number of citations steadily increasing with time. Further, this check does not indicate the extensive use of Latin hypercube sampling in analyses documented in the technical report literature. Thus, the use of Latin hypercube sampling is extensive and growing. As an indication of the interest in Latin hypercube sampling, the original article was recently declared a *Technometrics* classic in experimental design [216].

The growing use of Latin hypercube sampling and other techniques for the propagation and analysis of uncertainty derives from the recognition that it is not enough just to report the results of an analysis. For the analysis to be useful in a decision making context, it also necessary to assess and report how much confidence should be placed in the results of the analysis (e.g. see the recommendations given in quotes reproduced in Ref. [7]).

4. Comparison of random and Latin hypercube sampling

Because of its efficient stratification properties, Latin hypercube sampling is primarily intended for use with long-running models. When a model can be evaluated quickly, there is little reason to use Latin hypercube sampling. However, due to their computational complexity and expense, long-running models do not constitute convenient vehicles for comparing random and Latin hypercube sampling. For this reason, the present section will use two relatively simple functions (i.e. models) to compare random and Latin hypercube sampling. No comparisons with stratified sampling are made because the stratification used in a real analysis will always depend on the goals of the analysis and the properties of the model(s) used in the analysis. In particular, the efficacy of stratified sampling derives from an informed selection of strata of unequal probability.

4.1. Monotonic function

The function

$$f_1(U,V) = U + V + UV + U^2 + V^2 + U\min\{\exp(3V), 10\} \tag{4.1}$$

is monotonic for positive values of its arguments U, V and thus reasonably well behaved. For the purpose of comparing random and Latin hypercube sampling, U and V are assumed to be uncorrelated and uniformly disturbed on [1.0, 1.5] and [0, 1], respectively.

Both random and Latin hypercube sampling can be used to estimate the distribution of f that derives from the distributions assigned to U and V. To illustrate the robustness (i.e. stability) of results obtained with the two sampling procedures, 10 samples of size 25, 50 and 100 are generated for each procedure and the associated CDFs for f_1 constructed. The CDFs constructed for Latin hypercube sampling show less variability from sample to sample than the CDFs constructed for random sampling (Fig. 6). Thus, Latin hypercube sampling is producing a

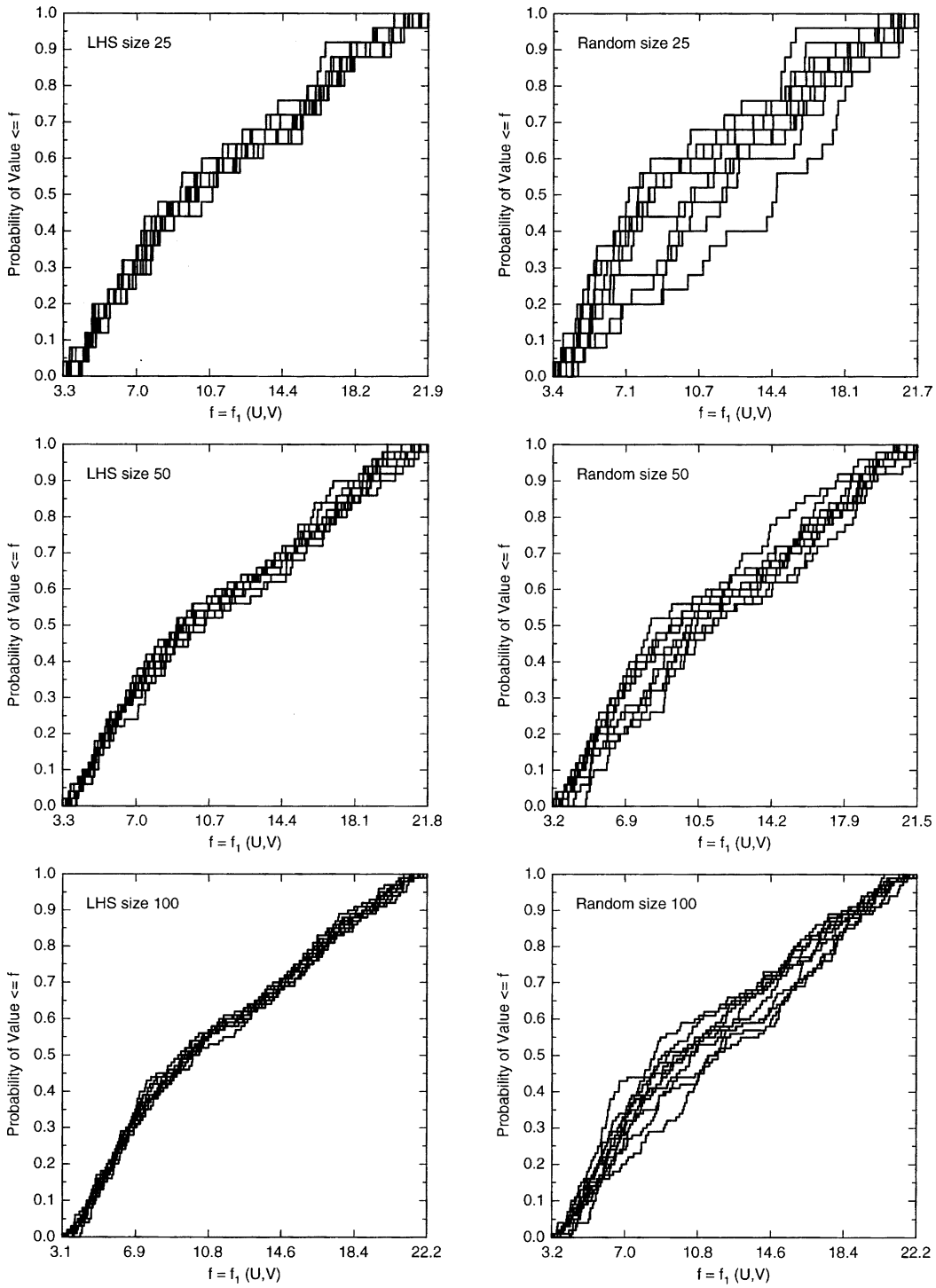

Fig. 6. Comparison of estimated CDFs for monotonic function $f_1(U, V)$ in Eq. (4.1) obtained with 10 replicated random and Latin hypercube samples of size 25, 50 and 100.

more stable estimate for the CDF than is being produced by random sampling, which is consistent with the result in Theorem 3.1.

4.2. Nonmonotonic function

Unlike the function f_1 in Eq. (4.1), the following function is monotonic for positive values of one argument (i.e. U) and nonmonotonic for positive values of the other argument (i.e. V):

$$f_2(U, V) = U + V + UV + U^2 + V^2 + Ug(V) \tag{4.2}$$

where

$$h(V) = (V - 11/43)^{-1} + (V - 22/43)^{-1} + (V - 33/43)^{-1}$$
$$g(V) = h(V) \text{ if } |h(V)| < 10$$

$g(V) = 10$ if $h(V) \geq 10$
$g(V) = -10$ if $h(V) \leq -10$.

For the purpose of comparing random and Latin hypercube sampling, U and V are again assumed to be uncorrelated and uniformly distributed on [1.0, 1.5] and [0, 1], respectively. Consideration of samples of size 25, 50 and 100 illustrates that Latin hypercube sampling produces more stable CDF estimates than produced by random sampling (Fig. 7).

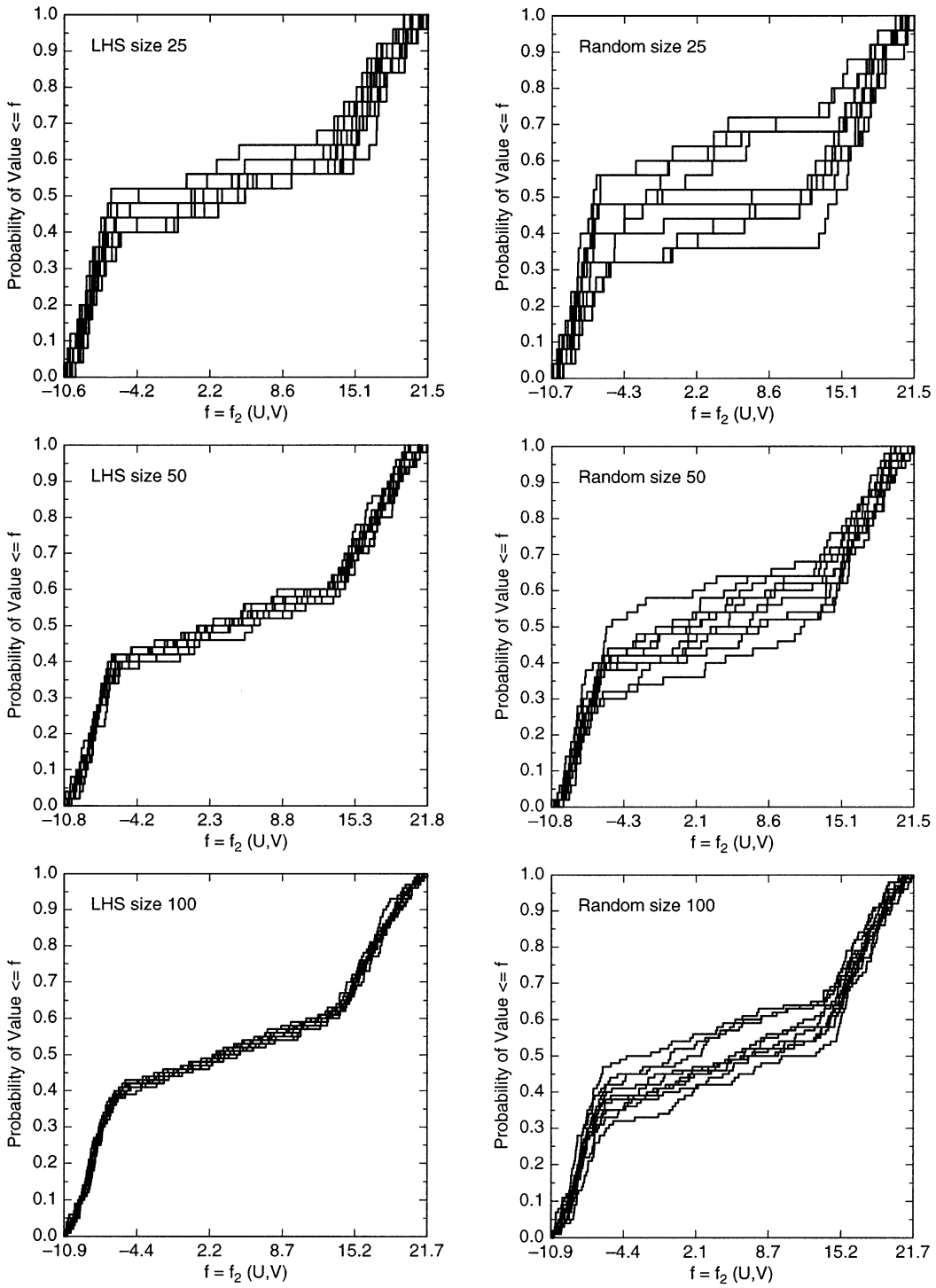

Fig. 7. Comparison of estimated CDFs for nonmonotonic function $f_2(U, V)$ in Eq. (4.2) obtained with 10 replicated random and Latin hypercube samples of size 25, 50 and 100.

5. Operations involving Latin hypercube sampling

5.1. Correlation control

As indicated in Eq. (1.2), the uncertainty in the inputs $x_1, x_2, \ldots, x_{nX}$ to an analysis can be represented by distributions $D_1, D_2, \ldots, D_{nX}$. If appropriate, correlations can also be specified between variables and form part of the definition of the corresponding probability space ($\mathscr{S}_{su}$, $\mathfrak{S}_{su}$, p_{su}). Given that $D_1, D_2, \ldots, D_{nX}$ are characterizing subjective uncertainty, correlations involving $x_1, x_2, \ldots, x_{nX}$ must in some sense derive from a belief that a particular value for one variable implies something about the possible values for one or more other variables (e.g. a low value for x_1 implies a high value for x_2, or a high value for x_3 implies a high value for x_5 and a low value for x_6), with the actual relationship being less strong than a strict functional dependence.

Two widely used possibilities exist for defining correlations between variables: the Pearson correlation coefficient (CC) and the Spearman rank correlation coefficient (RCC). For samples of the form in Eqs. (3.2) and (3.4), the CC between two variables, say x_j and x_k, is defined by

$$r_{x_j x_k} = \frac{\sum_{i=1}^{nS} (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\left[\sum_{i=1}^{nS} (x_{ij} - \bar{x}_j)^2\right]^{1/2}\left[\sum_{i=1}^{nS} (x_{ik} - \bar{x}_k)^2\right]^{1/2}}, \tag{5.1}$$

where

$$\bar{x}_j = \sum_{i=1}^{nS} x_{ij}/nS, \qquad \bar{x}_k = \sum_{i=1}^{nS} x_{ik}/nS.$$

The CC takes on values between -1 and 1 and provides a measure of the strength of the linear relationship between two variables, with variables tending to move in the same direction and in opposite directions for positive and negative CCs, respectively, and with gradations in the absolute value of the CC between 0 and 1 corresponding to a trend from no linear relationship to an exact linear relationship.

The RCC is defined similarly to the CC but with rank-transformed data. Specifically, the smallest value of a variable is given a rank of 1; the next largest value is given a rank of 2; and so on up to the largest value, which is given a rank equal to the sample size nS. In the event of ties, average ranks are assigned. The RCC is then calculated in the same manner as the CC except for the use of rank-transformed data. Specifically,

$$R_{x_j x_k} = \frac{\sum_{i=1}^{nS} [R(x_{ij}) - \bar{R}(x_j)][R(x_{ik}) - \bar{R}(x_k)]}{\left\{\sum_{i=1}^{nS} [R(x_{ij}) - \bar{R}(x_j)]^2\right\}^{1/2}\left\{\sum_{i=1}^{nS} [R(x_{ik}) - \bar{R}(x_k)]^2\right\}^{1/2}}, \tag{5.2}$$

where $R(x_{ij})$ and $R(x_{ik})$ denote the rank-transformed values of x_{ij} and x_{ik}, respectively, and $\bar{R}(x_j) = \bar{R}(x_k) = (nS+1)/2$. Like the CC, the RCC takes on values between -1 and 1 but provides a measure of the strength of the monotonic relationship between two variables.

In the authors' opinion, most individuals intuitively think in terms of RCCs rather than CCs when correlations are used in association with assessments of subjective uncertainty. In particular, what is usually possessed is some idea of the extent to which large and small values for one variable should be associated with large and small values for another variable. This is exactly the type of information that is quantitatively captured by RCCs. Therefore, this section will discuss the imposition of a rank correlation structure on random and LHSs.

An effective technique for imposing rank correlations has been proposed by Iman and Conover [151]. This technique has several desirable properties including (i) distribution independence in the sense that it can be applied to all types of distributions, (ii) simplicity in that no unusual mathematical techniques are required in its implementation, (iii) the stratification associated with Latin hypercube sampling is preserved, (iv) the marginal distributions for the individual sample variables are preserved, and (v) complex correlation structures involving many variables can be imposed on a sample.

The following discussion provides an overview of the Iman/Conover procedure for inducing a desired rank correlation structure on either a random or an LHS and is adapted from Section 3.2 of Helton [217]. The procedure begins with a sample of size m from the n input variables under consideration. This sample can be represented by the $m \times n$ matrix

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \tag{5.3}$$

where x_{ij} is the value for variable j in sample element i. Thus, the rows of $\mathbf{X}$ correspond to sample elements, and the columns of $\mathbf{X}$ contain the sampled values for individual variables.

The procedure is based on rearranging the values in the individual columns of $\mathbf{X}$ so that a desired rank correlation structure results between the individual variables. For convenience, let the desired correlation structure be represented by the $n \times n$ matrix

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \tag{5.4}$$

where c_{kl} is the desired rank correlation between variables x_k and x_l.

Although the procedure is based on rearranging the values in the individual columns of $\mathbf{X}$ to obtain a new matrix $\mathbf{X}^*$ that has a rank correlation structure close to that described by $\mathbf{C}$, it is not possible to work directly with $\mathbf{X}$. Rather, it is necessary to define a new matrix

$$\mathbf{S} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & & \vdots \\ s_{m1} & s_{m2} & \cdots & s_{mn} \end{bmatrix} \tag{5.5}$$

that has the same dimensions as $\mathbf{X}$, but is otherwise independent of $\mathbf{X}$. Each column of $\mathbf{S}$ contains a random permutation of the m van der Waerden scores $\Phi^{-1}(i/m+1)$, $i = 1, 2, \ldots, m$, where Φ^{-1} is the inverse of the standard normal distribution (Ref. [218], p. 317). The matrix $\mathbf{S}$ is then rearranged to obtain the correlation structure defined by $\mathbf{C}$. This rearrangement is based on the Cholesky factorization of $\mathbf{C}$ (Ref. [219], p. 89). That is, a lower triangular matrix $\mathbf{P}$ is constructed such that

$$\mathbf{C} = \mathbf{P}\mathbf{P}^{\mathrm{T}}. \tag{5.6}$$

This construction is possible because $\mathbf{C}$ is a symmetric, positive-definite matrix (Ref. [219], p. 88).

If the correlation matrix associated with $\mathbf{S}$ is the $n \times n$ identity matrix (i.e. if the correlations between the values in different columns of $\mathbf{S}$ are zero), then the correlation matrix for

$$\mathbf{S}^* = \mathbf{S}\mathbf{P}^{\mathrm{T}} \tag{5.7}$$

is $\mathbf{C}$ (Ref. [220], p. 25). At this point, the success of the procedure depends on the following two conditions: (i) that the correlation matrix associated with $\mathbf{S}$ be close to the $n \times n$ identity matrix, and (ii) that the correlation matrix for $\mathbf{S}^*$ be approximately equal to the rank correlation matrix for $\mathbf{S}^*$. If these two conditions hold, then the desired matrix $\mathbf{X}^*$ can be obtained by simply rearranging the values in the individual columns of $\mathbf{X}$ in the same rank order as the values in the individual columns of $\mathbf{S}^*$. This is the first time that the variable values contained in $\mathbf{X}$ enter into the correlation process. When $\mathbf{X}^*$ is constructed in this manner, it will have the same rank correlation matrix as $\mathbf{S}^*$. Thus, the rank correlation matrix for $\mathbf{X}^*$ will approximate $\mathbf{C}$ to the same extent that the rank correlation matrix for $\mathbf{S}^*$ does.

The condition that the correlation matrix associated with $\mathbf{S}$ be close to the identity matrix is now considered. For convenience, the correlation matrix for $\mathbf{S}$ will be represented by $\mathbf{E}$. Unfortunately, $\mathbf{E}$ will not always be the identity matrix. However, it is possible to make a correction for this. The starting point for this correction is the Cholesky factorization for $\mathbf{E}$:

$$\mathbf{E} = \mathbf{Q}\mathbf{Q}^{\mathrm{T}}. \tag{5.8}$$

This factorization exists because $\mathbf{E}$ is a symmetric, positive-definite matrix. The matrix $\mathbf{S}^*$ defined by

$$\mathbf{S}^* = \mathbf{S}(\mathbf{Q}^{-1})^{\mathrm{T}}\mathbf{P}^{\mathrm{T}} \tag{5.9}$$

has $\mathbf{C}$ as its correlation matrix. In essence, multiplication of $\mathbf{S}$ by $(\mathbf{Q}^{-1})^{\mathrm{T}}$ transforms $\mathbf{S}$ into a matrix whose associated correlation matrix is the $n \times n$ identity matrix; then, multiplication by $\mathbf{P}^{\mathrm{T}}$ produces a matrix whose associated correlation matrix is $\mathbf{C}$. As it is not possible to be sure that $\mathbf{E}$ will be an identity matrix, the matrix $\mathbf{S}^*$ used in the procedure to produce correlated input should be defined in the corrected form shown in Eq. (5.9) rather than in the uncorrected form shown in Eq. (5.7).

The condition that the correlation matrix for $\mathbf{S}^*$ be approximately equal to the rank correlation matrix for $\mathbf{S}^*$ depends on the choice of the scores used in the definition of $\mathbf{S}$. On the basis of empirical investigations, Iman and Conover [151] found that van der Waerden scores provided an effective means of defining $\mathbf{S}$, and these scores are incorporated into the rank correlation procedure in the widely used LHS program [165]. Other possibilities for defining these scores exist, but have not been extensively investigated. The user should examine the rank correlation matrix associated with $\mathbf{S}^*$ to ensure that it is close to the target correlation matrix $\mathbf{C}$. If this is not the case, the construction procedure used to obtain $\mathbf{S}^*$ can be repeated until a suitable approximation to $\mathbf{C}$ is obtained. Results given in Iman and Conover [151] indicate that the use of van der Waerden scores leads to rank correlation matrices for $\mathbf{S}^*$ that are close to the target matrix $\mathbf{C}$.

As a single example, the effects of imposing rank correlations of 0.00, 0.25, 0.50, 0.75, 0.90 and 0.99 on a pair of variables are shown in Fig. 8. The results of various rank-correlation assumptions with a variety of marginal distributions are illustrated by Iman and Davenport [221,222].

The control of orthogonality and the induction of correlations within LHSs are areas of much research interest, and a number of results exist in this area in addition to the original Iman and Conover rank correlation techniques discussed in this section [223–241].

5.2. Reweighting of samples

Once a sampling-based uncertainty study has been performed, it is sometimes necessary to assess the effects that arise from changed definitions for the distributions $D_1, D_2, \ldots, D_{nX}$ in Eq. (1.2). If the model under consideration is expensive to evaluate, it is desirable to perform this assessment without reevaluating (i.e. rerunning) the model. When the distributions but not the ranges of the variables change, this assessment can be carried out with a reweighting technique developed by Iman and Conover [242].

Latin hypercube sampling as described in Section 3.1 is based on dividing the range of each variable into nS

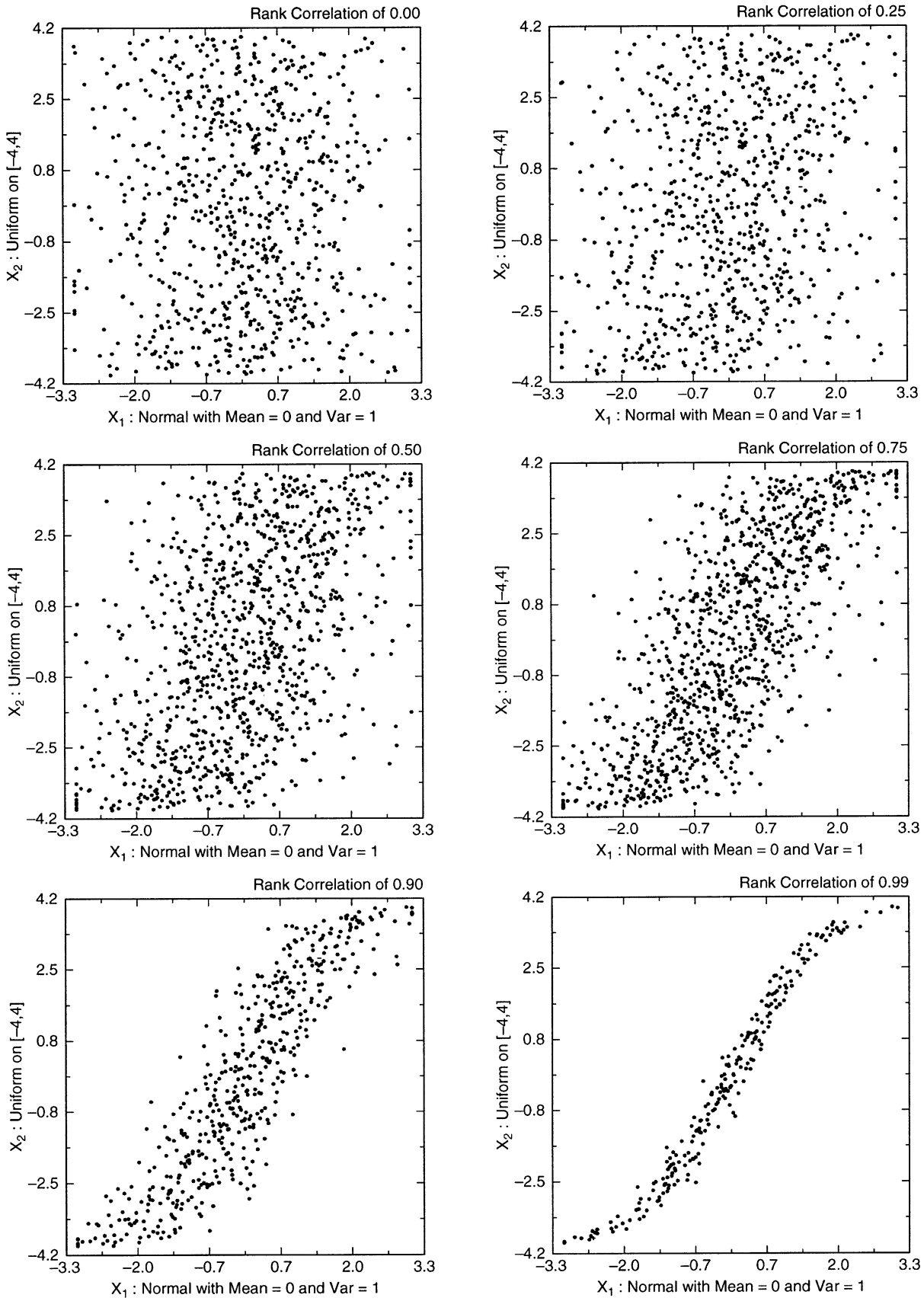

Fig. 8. Examples of rank correlations of 0.00, 0.25, 0.50, 0.75, 0.90 and 0.99 imposed with the Iman/Conover restricted pairing technique for an LHS of size $nS = 1000$.

intervals of equal probability, where nS is the sample size. The Iman/Conover reweighting technique is based on a generalization of Latin hypercube sampling that involves the division of variable ranges into intervals of unequal probability.

For this generalization of an LHS size nS from the variables $x_1, x_2, \ldots, x_{nX}$, the range of each variable x_j is divided into nS mutually exclusive intervals I_{ij}, $i = 1, 2, \ldots, nS$, and one value x_{ij}, $i = 1, 2, \ldots, nS$, of x_j is randomly selected from each interval I_{ij}. The preceding variable values (i.e. x_{ij}, $i = 1, 2, \ldots, nS$, $j = 1, 2, \ldots, nX$) are now used as described in Section 3.1 to generate an LHS. Specifically, the nS values for x_1 are randomly paired without replacement with the nS values for x_2. The resultant nS pairs are randomly combined without replacement with the nS values for x_3 to produce nS triples. This process is continued until nS nX-tuples are produced, with these nX-tuples constituting the LHS

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{i,nX}], \quad i = 1, 2, \ldots, nS. \tag{5.10}$$

The preceding division of the ranges of the variables into the intervals I_{ij} produces a corresponding division of $\mathscr{S}_{su}$ into nS^{nX} cells. Specifically, each cell is of the form

$$\mathscr{C}_{\mathbf{n}} = I_{k1} \times I_{l2} \times \cdots \times I_{m,nX}, \tag{5.11}$$

where $\mathbf{n} = [k, l, \ldots, m]$ is a vector of nX integers between 1 and nS that designates one of the nS^{nX} mutually exclusive cells into which $\mathscr{S}_{su}$ has been partitioned. Further, the probability $\text{prob}(\mathscr{C}_{\mathbf{n}})$ of $\mathscr{C}_{\mathbf{n}}$ can be calculated from the definition of $(\mathscr{S}_{su}, \mathfrak{S}_{su}, p_{su})$. For example,

$$\text{prob}(\mathscr{C}_{\mathbf{n}}) = \text{prob}(I_{k1})\text{prob}(I_{l2}) \cdots \text{prob}(I_{m,nX}) \tag{5.12}$$

if the x_js are independent.

Theorem 5.1. *If* $\mathbf{x}_i$, $i = 1, 2, \ldots, nS$, *is an LHS of the form indicated in Eq.* (5.10), $\mathscr{C}_{\mathbf{n}_i}$, $i = 1, 2, \ldots, nS$, *designates the cell in Eq.* (5.11) *that contains* $\mathbf{x}_i$, $\mathbf{f}$ *is the function in Eq.* (1.1), *and* g *is an arbitrary function, then*

$$T = \sum_{i=1}^{nS} nS^{nX-1} \, \text{prob}(\mathscr{C}_{\mathbf{n}_i}) g[\mathbf{f}(\mathbf{x}_i)] \tag{5.13}$$

is an unbiased estimator of the expected value of $g[\mathbf{f}(\mathbf{x})]$ (*Theorem 1, p. 1760, Ref.* [242]).

The preceding result reduces to the unbiasedness of the estimator in Eq. (3.5) when Latin hypercube sampling with equal probability intervals is used (i.e. $\text{prob}(\mathscr{C}_{\mathbf{n}_i}) = 1/nS^{nX}$) and $\mathbf{f}(\mathbf{x}_i)$ is real valued (i.e. $y_i = \mathbf{f}(\mathbf{x}_i)$). The importance of Theorem 5.1 is that it allows a recalculation of expected values, moments and distribution functions that result from changed distribution assumptions without a rerunning of the model under consideration. Specifically, the same values for $g[\mathbf{f}(\mathbf{x}_i)]$ are used in conjunction with new values for $\text{prob}(\mathscr{C}_{\mathbf{n}_i})$ calculated for the changed distributions for the elements of $\mathbf{x}$. A related result is given by Beckman and McKay [243].

5.3. Replication of samples

A brief overview of the variability in statistics obtained with Latin hypercube sampling is given in Section 3.2. The variability results when the same quantity is repeatedly estimated with independently generated samples of the same size. In essence, this variability is a measure of the numerical error in using a sampling-based (i.e. Monte Carlo) procedure in the estimation of an integral. Unfortunately, the theoretical results indicated in Section 3.2 do not

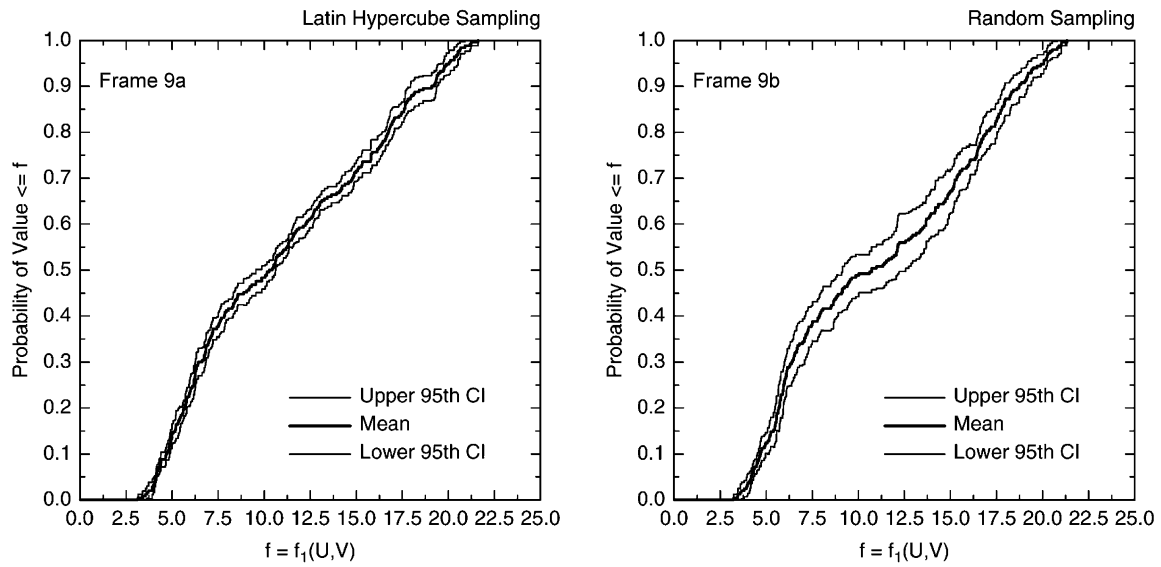

Fig. 9. Upper and lower bounds on 0.95 confidence intervals (CIs) for cumulative probabilities associated with function $f_1(U, V)$ in Eq. (4.1) obtained from $nR = 10$ samples of size $nS = 25$ each (see Fig. 6 for CDFs): (a) Latin hypercube sampling, and (b) random sampling.

lead in any convenient way to error estimates in real analyses.

In practice, a replicated sampling procedure proposed by Iman [244] provides a more effective approach to estimating the potential sampling error in quantities derived from Latin hypercube sampling. With this procedure, the LHS in Eq. (3.4) is repeatedly generated with different random seeds. These samples are used to produce a sequence of values T_r, $r = 1, 2, \ldots, nR$, for the statistic T in Eq. (3.5), where nR is the number of replicated samples. Then,

$$\bar{T} = \sum_{r=1}^{nR} T_r/nR \tag{5.14}$$

and

$$\mathrm{SE}(\bar{T}) = \left[\sum_{r=1}^{nR} (T_r - \bar{T})^2/nR(nR-1)\right]^{1/2} \tag{5.15}$$

provide an additional estimate for T and an estimate of the standard error for this estimate of T. The t-distribution with $nR - 1$ degrees of freedom can be used to obtain a confidence interval for the estimate for $\bar{T}$. Specifically, the $1 - \alpha$ confidence interval is given by $\bar{T} \pm t_{1-\alpha/2}\mathrm{SE}(\bar{T})$, where $t_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the t-distribution with $nR - 1$ degrees of freedom (e.g. $t_{1-\alpha/2} = 2.262$ for $\alpha = 0.05$ and $nR = 10$; Ref. [218], Table A25).

As an example, 0.95 confidence intervals for the cumulative probabilities associated with individual values in the range of the function f_1 defined in Eq. (4.1) are shown in Fig. 9, with the 10 replicated LHSs producing narrower confidence intervals than the 10 random samples. The confidence intervals in Fig. 9 were calculated for individual values on the abscissa and then connected to obtain the confidence-interval curves (i.e. the curves of upper and lower bounds). Thus, the confidence intervals apply to individual cumulative probabilities rather than to an entire CDF.

6. Example uncertainty and sensitivity analysis

An example uncertainty and sensitivity analysis involving a model for two-phase fluid flow follows. The analysis problem is briefly described (Section 6.1), and then techniques for the presentation of uncertainty analysis results are described and illustrated (Section 6.2). The section then concludes with illustrations of various sensitivity analysis procedures, including examination of scatterplots (Section 6.3), regression-based techniques (Section 6.4), and searches for nonrandom patterns (Section 6.5).

6.1. Analysis problem

The following examples use results from an uncertainty and sensitivity analysis of a large model for two-phase fluid

Table 1
Example elements of $\mathbf{x}_{su}$ in the 1996 WIPP PA (see Table 5.1, Ref. [245], Table 1, Ref. [250] and App. PAR, Ref. [190] for complete listings of the $nV = 57$ elements of $\mathbf{x}_{su}$ and sources of additional information)

ANHBCEXP—Brooks–Corey pore distribution parameter for anhydrite (dimensionless). Distribution: Student's with 5 degrees of freedom. Range: 0.491–0.842. Mean, Median: 0.644

ANHBCVGP—Pointer variable for selection of relative permeability model for use in anhydrite. Distribution: Discrete with 60% 0, 40% 1. Value of 0 implies Brooks–Corey model; value of 1 implies van Genuchten–Parker model

ANHCOMP—Bulk compressibility of anhydrite (Pa^{-1}). Distribution: Student's with 3 degrees of freedom. Range: 1.09×10^{-11} to 2.75×10^{-10} Pa^{-1}. Mean, Median: 8.26×10^{-11} Pa^{-1}. Correlation: −0.99 rank correlation with ANHPRM

ANHPRM—Logarithm of intrinsic anhydrite permeability (m^2). Distribution: Student's with 5 degrees of freedom (see Fig. 10). Range: −21.0 to −17.1 (i.e. permeability range is 1×10^{-21} to $1 \times 10^{-17.1}$ m^2). Mean, Median: −18.9. Correlation: −0.99 rank correlation with ANHCOMP

BHPRM—Logarithm of intrinsic borehole permeability (m^2). Distribution: Uniform. Range: −14 to −11 (i.e. permeability range is 1×10^{-14} to 1×10^{-11} m^2). Mean, median: −12.5.

BPCOMP—Logarithm of bulk compressibility of brine pocket (Pa^{-1}). Distribution: Triangular. Range: −11.3 to −8.00 (i.e. bulk compressibility range is $1 \times 10^{-11.3}$ to 1×10^{-8} Pa^{-1}). Mean, mode: −9.80, −10.0. Correlation: −0.75 rank correlation with BPPRM

BPPRM—Logarithm of intrinsic brine pocket permeability (m^2). Distribution: Triangular. Range: −14.7 to −9.80 (i.e. permeability range is $1 \times 10^{-14.7}$ to $1 \times 10^{-9.80}$ m^2). Mean, mode: −12.1, −11.8. Correlation: −0.75 with BPCOMP

HALCOMP—Bulk compressibility of halite (Pa^{-1}). Distribution: Uniform. Range: 2.94×10^{-12} to 1.92×10^{-10} Pa^{-1}. Mean, median: 9.75×10^{-11} Pa^{-1}. Correlation: −0.99 rank correlation with HALPRM

HALPOR—Halite porosity (dimensionless). Distribution: Piecewise uniform (see Fig. 10). Range: 1.0×10^{-3} to 3×10^{-2}. Mean, median: 1.28×10^{-2}, 1.00×10^{-2}

HALPRM—Logarithm of halite permeability (m^2). Distribution: Uniform. Range: −24 to −21 (i.e. permeability range is 1×10^{-24} to 1×10^{-21} m^2). Mean, median: −22.5. Correlation: −0.99 rank correlation with HALCOMP

SALPRES—Initial brine pressure, without the repository being present, at a reference point located in the center of the combined shafts at the elevation of the midpoint of MB 139 (Pa). Distribution: Uniform. Range: 1.104×10^7 to 1.389×10^7 Pa. Mean, median: 1.247×10^7 Pa

SHRBRSAT—Residual brine saturation in shaft (dimensionless). Distribution: Uniform. Range: 0–0.4. Mean, median: 0.2

SHRGSSAT—Residual gas saturation in shaft (dimensionless). Distribution: Uniform. Range: 0–0.4. Mean, median: 0.2

WASTWICK—Increase in brine saturation of waste due to capillary forces (dimensionless). Distribution: Uniform. Range: 0–1. Mean, median: 0.5

WGRCOR—Corrosion rate for steel under inundated conditions in the absence of CO_2 (m/s). Distribution: Uniform. Range: $0–1.58 \times 10^{-14}$ m/s. Mean, median: 7.94×10^{-15} m/s

WGRMICI—Microbial degradation rate for cellulose under inundated conditions (mol/kg s). Distribution: Uniform. Range: 3.17×10^{-10} to 9.51×10^{-9} mol/kg s. Mean, median: 4.92×10^{-9} mol/kg s.

WMICDFLG—Pointer variable for microbial degradation of cellulose. Distribution: Discrete, with 50% 0, 25% 1, 25% 2. WMICDFLG = 0, 1, 2 implies no microbial degradation of cellulose, microbial degradation of only cellulose, microbial degradation of cellulose, plastic and rubber

WPRTDIAM—Waste particle diameter (m). Distribution: Loguniform. Range: 4.0×10^{-5} to 2.0×10^{-1} m. Mean, median: 2.83×10^{-3} m

WRGSSAT—Residual gas saturation in waste (dimensionless). Distribution: Uniform. Range: 0–0.15. Mean, median: 0.075

WTAUFAIL—Shear strength of waste (Pa). Distribution: Uniform. Range: 0.05–10 Pa. Mean, median: 5.03 Pa

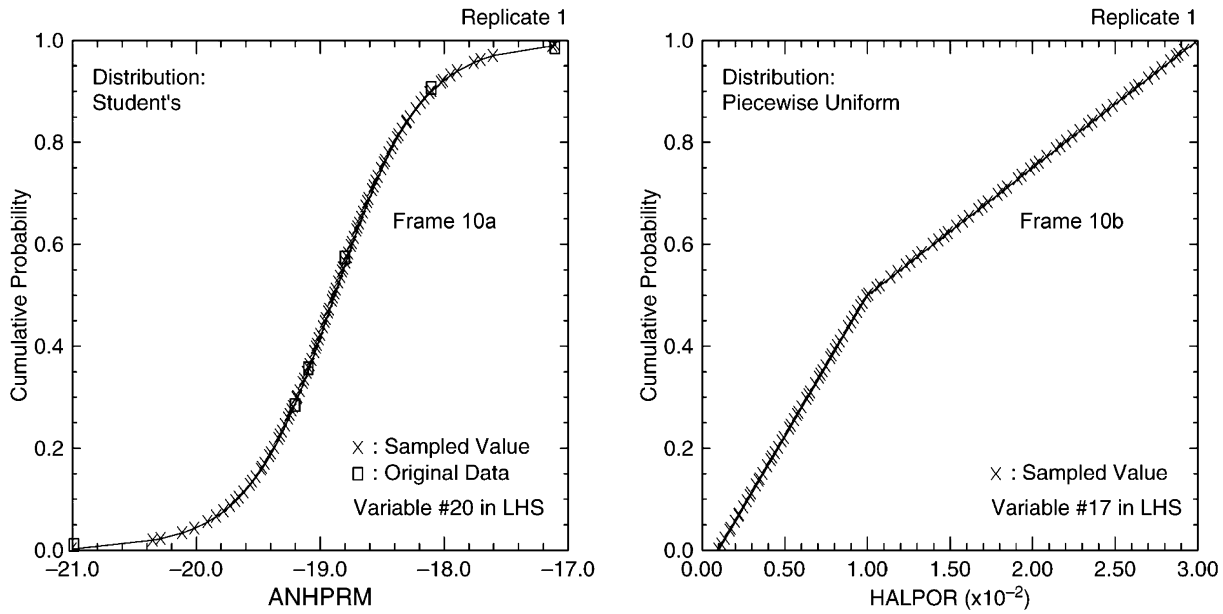

Fig. 10. Distributions used to characterize subjective uncertainty in ANHPRM and HALPOR in 1996 WIPP PA (see Appendix, Ref. [245], for distributions assigned to all uncertain variables included in 1996 WIPP PA).

flow (Ref. [245], Chapts. 7,8; Refs. [246–248]) carried out in support of the 1996 CCA for the WIPP, which is being developed by the DOE for the geologic (i.e. deep underground) disposal of transuranic radioactive waste [190,249]. The indicated model involves the numerical solution of a system of nonlinear partial differential equations and is implemented by the BRAGFLO program (Section 4.2, Ref. [245]; Ref. [246]).

The 1996 WIPP PA considered $nX = 57$ uncertain inputs (Table 1), of which 31 were used in the two-phase flow analysis and 26 were used in other parts of the PA (see Section 7). The distributions assigned to these variables (Fig. 10) are intended to characterize subjective uncertainty, correspond to the distributions in Eq. (1.2), and define the probability space $(\mathscr{S}_{su}, \Xi_{su}, p_{su})$.

Latin hypercube sampling was used to generate $nR = 3$ replicated samples of size $nS = 100$ each (Section 5.3) for a total of 300 sample elements. For convenience, these replicates are referred to as R1, R2 and R3, respectively. The Iman/Conover restricted pairing technique (Section 5.1) was used to induce specified rank correlations for three pairs of variables (Table 1) and to keep correlations between all other variables close to zero (Table 2).

As is typical of most studies of real systems, the original analysis involved a large number of dependent variables, of which only 11 will be used for illustration in this section (Table 3). The variables in Table 3 were calculated for three distinct sets of conditions designated by E0, E1 and E2 in the 1996 WIPP PA, where E0 corresponds to undisturbed conditions (i.e. no human disruption of the repository), E1 corresponds to a single drilling intrusion through the repository that penetrates an area of pressurized brine in a geologic formation beneath the repository, and E2 corresponds to a single drilling intrusion through the repository that does not penetrate pressurized brine.

6.2. Uncertainty analysis

In this example, the model predictions are functions rather than single numbers as indicated in conjunction with Fig. 1. The distributions of curves in Fig. 11 constitute one way of displaying the uncertainty in these functions that

Table 2
Example rank correlations in the LHS that constitutes replicate R1 in the 1996 WIPP PA

	WGRCOR	WMICDFLG	HALCOMP	HALPRM	ANHCOMP	ANHPRM	BPCOMP	BPPRM
WGRCOR	1.0000							
WMICDFLG	0.0198	1.0000						
HALCOMP	0.0011	0.0235	1.0000					
HALPRM	−0.0068	−0.0212	−0.9879	1.0000				
ANHCOMP	0.0080	0.0336	−0.0123	−0.0025	1.0000			
ANHPRM	0.0049	−0.0183	0.0037	0.0113	−0.9827	1.0000		
BPCOMP	0.0242	0.1071	−0.0121	0.0057	−0.0184	0.0078	1.0000	
BPPRM	−0.0514	−0.0342	0.0035	0.0097	0.0283	−0.0202	−0.7401	1.0000

Table 3
Predicted variables (i.e. elements of **y** in Eq. (1.1)) used to illustrate uncertainty and sensitivity analysis results for two-phase fluid flow model (see Table 7.1.1, Ref. [245], for additional information)

BRAALIC	Cumulative brine flow (m^3) from anhydrite marker beds (AMBs) into disturbed rock zone (DRZ) surrounding repository (i.e. BRAABNIC + BRAABSIC + BRM38NIC + BRM38SIC + BRM39NIC + BRM39SIC)
BRAABNIC	Cumulative brine flow (m^3) out of anhydrite marker beds A and B into north end of DRZ
BRAABSIC	Same as BRAABNIC but into south end of DRZ
BRM38NIC	Cumulative brine flow (m^3) out of anhydrite marker bed 138 into north end of DRZ
BRM38SIC	Same as BRM38NIC but into south end of DRZ
BRM39NIC	Cumulative brine flow (m^3) out of anhydrite marker bed 139 into north end of DRZ
BRM39SIC	Same as BRM39NIC but into south end of DRZ
BRNREPTC	Cumulative brine flow (m^3) into repository from all sources
GAS_MOLE	Cumulative gas production (mole) in repository due to corrosion of iron and microbial degradation of cellulose
PORVOL_T	Total pore volume (m^3) in repository
WAS_SATB	Brine saturation (dimensionless) in lower waste panel (i.e. the southern waste panel, which in the numerical implementation of the analysis is the waste panel that is penetrated by a drilling intrusion for the E1 and E2 scenarios)

results from uncertainty in model input. However, the model predictions at individual times are real valued and thus can be displayed as CDFs or CCDFs. A popular presentation format [251] is to display estimates for the CDF, the corresponding density function, and the mean in a single plot (Fig. 12).

For distributions of curves such as those in Fig. 11, summaries can be obtained by plotting mean and percentile values of the dependent variable for individual values on the abscissa (Fig. 13). Conceptually, a vertical line is drawn through a point on the abscissa and the curves above this point. If a sample of size nS is involved, this results in selecting nS values for the dependent variable (i.e. the nS values above the point on the abscissa). These values can then be used to estimate a mean, a median, and various percentiles. Connecting these estimates for a sequence of values on the abscissa produces summary plots of the form shown in Fig. 13.

The purpose of replicating the LHS in this example was to obtain an indication of the stability of the resultant distribution estimates with an LHS of size 100. In this analysis, these estimates were quite stable (e.g. Fig. 14). Similar stability has also been observed in other studies [32,171,252,253].

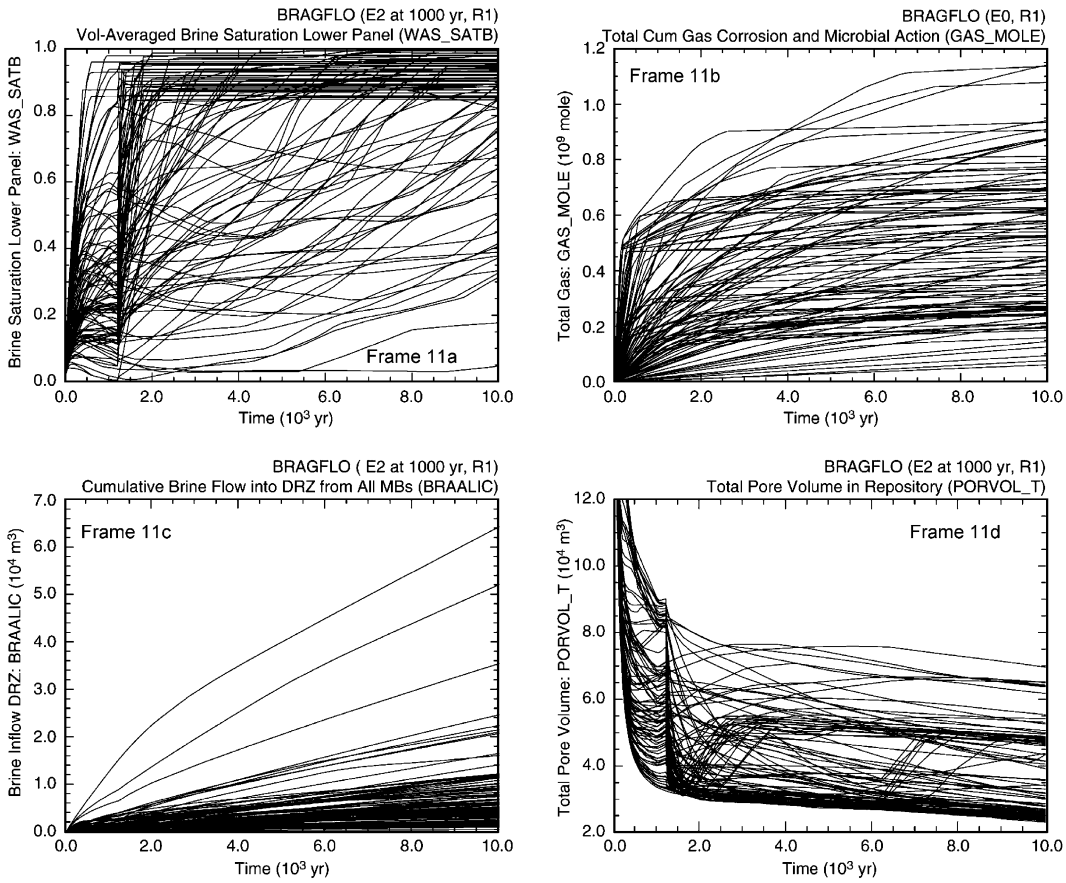

Fig. 11. Time-dependent results used to illustrate sensitivity analysis techniques: (a) saturation in lower waste panel with an E2 intrusion at 1000 yr (E2:WAS_SATB), (b) total cumulative gas generation due to corrosion and microbial degradation of cellulose under undisturbed (i.e. E0) conditions (E0:GAS_MOLE), (c) cumulative brine flow into disturbed rock zone (DRZ) surrounding repository with an E2 intrusion at 1000 yr (E2:BRAALIC); and (d) total pore volume in repository with an E2 intrusion at 1000 yr (E2:PORVOL_T).

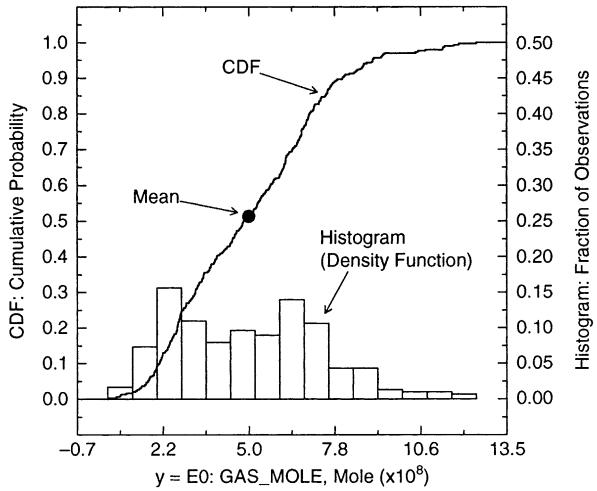

Fig. 12. Presentation of estimated CDF, mean, and density function for $y = $ E0:GAS_MOLE at 10,000 yr.

Presentation of multiple plots of the form shown in Fig. 12 can be cumbersome when a large number of predicted variables is involved. When these variables have the same units, box plots provide a way to present a compact summary of multiple distributions (Fig. 15). In this summary, the endpoints of the boxes are formed by the lower and upper quartiles of the data, that is, $x_{0.25}$ and $x_{0.75}$. The vertical line within the box represents the median, $x_{0.50}$. The mean is identified by the large dot. The bar on the right of the box extends to the minimum of $x_{0.75} + 1.5(x_{0.75} - x_{0.25})$ and the maximum value. In a similar manner, the bar on the left of the box extends to the maximum of $x_{0.25} - 1.5(x_{0.75} - x_{0.25})$ and the minimum value. The observations falling outside of these bars are shown in crosses. The flattened shape of box plots makes it possible to summarize multiple distributions in a small area and also facilitates comparisons of these distributions.

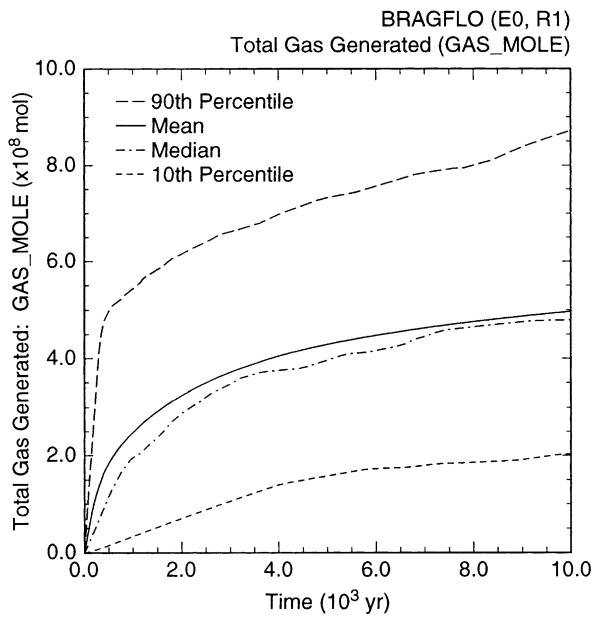

Fig. 13. Mean and percentile curves for $y = $ E0:GAS_MOLE for replicate R1.

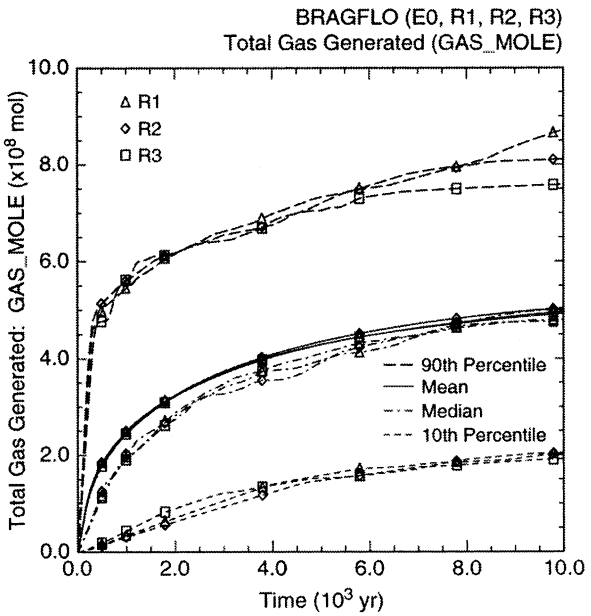

Fig. 14. Individual mean and percentile curves for $y = $ E0:GAS_MOLE for replicates R1, R2 and R3.

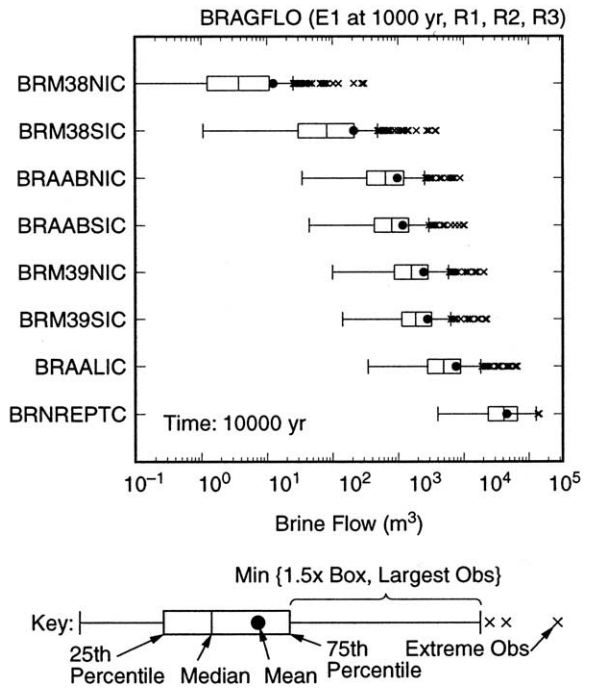

Fig. 15. Use of box plots to summarize cumulative brine flows over 10,000 yr in the vicinity of the repository for an E1 intrusion at 1000 yr into lower waste panel (see Table 3 for a description of individual variables).

6.3. Examination of scatterplots

The simplest sensitivity analysis procedure is an examination of the scatterplots associated with individual sampled variables and the particular model prediction under consideration (see Eq. (2.7)). If a variable has a substantial effect on the model prediction, then this will result in a discernible pattern in the corresponding scatterplot (Fig. 16); in contrast, little or no pattern will appear in the scatterplot in the absence of an effect. Further, the examination of multiple scatterplots can reveal interactions in the effects of variables. For example, large values of WAS_SATB tend to be associated with large values of BHPRM (Fig. 16a); however, given the occurrence of a large value for BHPRM, the resultant value for WAS_SATB is determined primarily by WRGSSAT (Fig. 16b). Latin hypercube sampling is a particularly effective procedure for the generation of scatterplots due to its full stratification across the range of each sampled variable.

6.4. Regression-based techniques

A more sophisticated approach to sensitivity analysis is to use formal search procedures to identify specific patterns in the mapping in Eq. (2.3). For example, regression-based techniques are often effective in identifying linear relationships and relationships that can be made linear by a suitable transformation (Ref. [217], Section 3.5). Stepwise regression analysis provides an efficient and informative way to carry out a regression-based sensitivity analysis, with variable importance being indicated by the order in which variables are selected in the stepwise procedure, the changes in R^2 values that occur as individual variables are added to the regression model, and the size of the SRCs for the variables included in the regression model. When the relationships between the sampled and predicted variables are nonlinear but monotonic, the rank transformation [254] is often effective in linearizing the underlying relationships and thus facilitating the use of regression-based techniques.

As an example, stepwise regression analyses for $y =$ E0:GAS_MOLE and $y =$ E2:BRAALIC with raw and rank-transformed data are presented in Table 4. For E0:GAS_MOLE, similar results are obtained with raw and rank-transformed data (i.e. the same variables are selected in both analyses and the final regression models have R^2 values of 0.85 and 0.82, respectively). For E2:BRAALIC, the use of rank-transformed data considerably improves the resolution of the analysis and produces a final regression model with six variables and an R^2 value of 0.90; in contrast, the use of raw data produces a final regression model with three variables and an R^2 value of 0.62.

An alternative to regression analysis is to calculate CCs or partial correlation coefficients (PCCs) between sampled and predicted variables (Ref. [217], Section 3.5). As with regression analysis, these coefficients can be calculated with raw or rank-transformed data, with the latter case producing RCCs and partial rank correlation coefficients (PRCCs). When the variables within the sample are independent (i.e. orthogonal), CCs and SRCs are equal, as is also the case for RCCs and standardized rank regression coefficients (SRRCs). Similar, but not entirely equivalent, measures of variable importance are given by SRCs and PCCs. Specifically, SRCs characterize the effect on the output

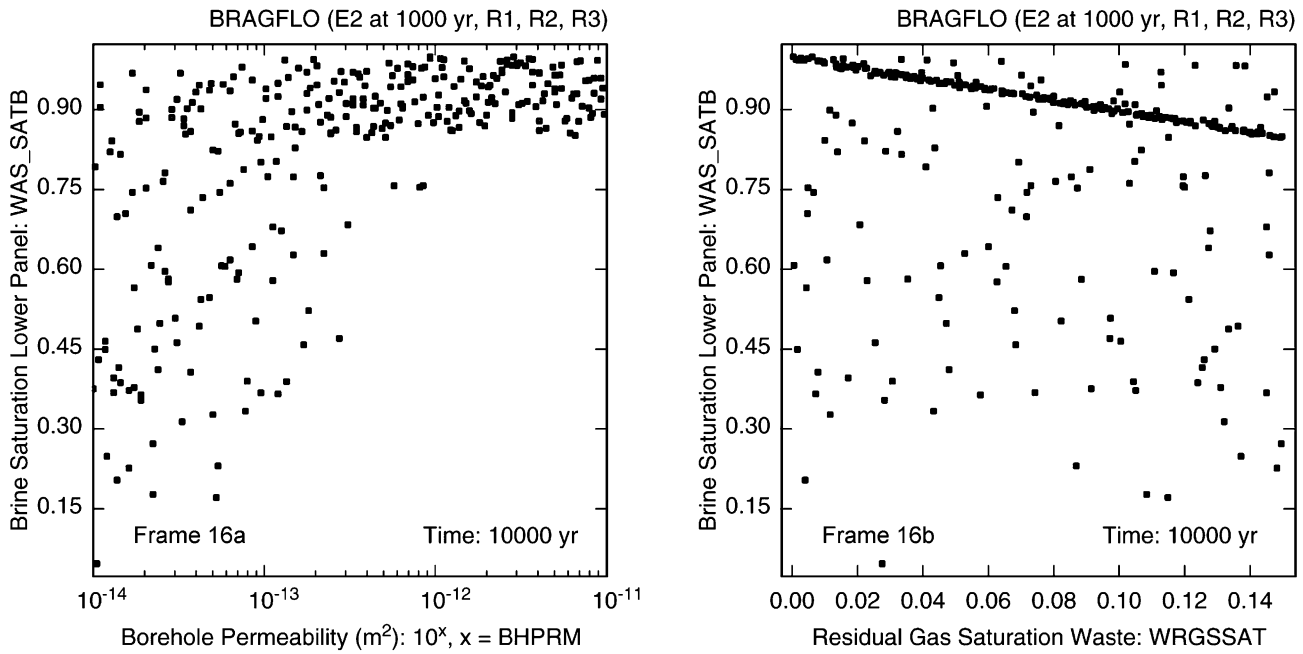

Fig. 16. Scatterplots for brine saturation in lower waste panel (WAS_SATB) at 10,000 yr for an E2 intrusion at 1000 yr into lower waste panel versus BHPRM and WRGSSAT.

variable that results from perturbing an input variable by a fixed fraction of its standard deviation, and PCCs characterize the strength of the linear relationship between an input and output variable after a correction has been made for the linear effects of the other input variables. Similar interpretations apply to SRRCs and PRCCs for rank-transformed variables. Although SRCs and PCCs are not equal, use of their absolute values to order variable importance produces identical importance orderings when the values for the individual variables within the sample are independent, as is also the case for SRRCs and PRCCs.

As in Fig. 11, model predictions are often time dependent. When this is the case, presenting stepwise regression analyses at multiple times in the format used in Table 4 can become quite unwieldy. In such situations, a more compact alternative is to present plots of time-dependent coefficients (Fig. 17). In particular, the coefficients are calculated at multiple times and then the coefficients for individual variables are connected to obtain the curves in Fig. 17. This presentation format is relatively compact and also displays how variable importance changes with time.

6.5. Searches for nonrandom patterns

Regression-based techniques are not always successful in identifying the relationships between sampled variables and model predictions. As an example, the regression analyses with raw and rank-transformed data in Table 5 perform poorly, with the final regression models having R^2 values of 0.33 and 0.20. Given the low R^2 values, there is little reason to believe that the variable orderings are meaningful or even that all the influential variables have been identified.

When regression-based approaches to sensitivity analysis do not yield satisfactory insights, important variables can be searched for by attempting to identify patterns in the mapping in Eq. (2.3) with techniques that are not predicated on searches for linear or monotonic relationships. Possibilities include use of (i) the F-statistic to identify changes in the mean value of y across the range of individual x_js, (ii) the χ^2-statistic to identify changes in the median value of y across the range of individual x_js, (iii) the Kruskal–Wallis statistic to identify changes in the distribution of y across the range of individual x_js, and (iv) the χ^2-statistic to identify nonrandom joint distributions involving y and individual x_js [255]. For convenience, the preceding will be referred to as tests for (i) common means (CMNs), (ii) common medians (CMDs), (iii) common locations (CLs), and (iv) statistical independence (SI), respectively.

The preceding statistics are based on dividing the values of x_j in Eq. (2.7) into intervals (Fig. 18). Typically, these intervals contain equal numbers of values for x_j (i.e. the intervals are of equal probability); however, this is not always the case (e.g. when x_j has a finite number of values of unequal probability). The calculation of the F-statistic for CMNs and the Kruskal–Wallis statistic for CLs involves only the division of x_j into intervals.

Table 4
Stepwise regression analyses with raw and rank-transformed data with pooled results from replicates R1, R2 and R3 (i.e. for a total of 300 observations) for output variables E0:GAS_MOLE and E2:BRAALIC at 10,000 yr

Step[a]	Raw data: y = E0:GAS_MOLE			Rank-transformed data: y = E0:GAS_MOLE		
	Variable[b]	SRC[c]	R^{2d}	Variable[b]	SRRC[e]	R^{2d}
1	WMICDFLG	0.65	0.41	WMICDFLG	0.62	0.39
2	HALPOR	0.59	0.76	HALPOR	0.57	0.72
3	WGRCOR	0.27	0.84	WGRCOR	0.28	0.80
4	WASTWICK	0.07	0.84	ANHPRM	0.08	0.81
5	ANHPRM	0.07	0.85	WASTWICK	0.07	0.81
6	SHRGSSAT	0.07	0.85	SHRGSSAT	0.07	0.82
	Raw data: y = E2:BRAALIC			Rank-transformed data: y = E2:BRAALIC		
	Variable	SRC	R^2	Variable	SRRC	R^2
1	ANHPRM	0.77	0.59	ANHPRM	0.91	0.83
2	WMICDFLG	−0.14	0.61	WMICDFLG	−0.15	0.85
3	SALPRES	0.09	0.62	BHPRM	0.13	0.87
4				HALPRM	0.12	0.88
5				SALPRES	0.10	0.89
6				WGRCOR	−0.05	0.90

[a] Steps in stepwise regression analysis with significance levels of $\alpha = 0.02$ and $\alpha = 0.05$ required of a variable for entry into and retention in a regression model, respectively.
[b] Variables listed in order of selection in regression analysis with ANHCOMP and HALCOMP excluded from entry into regression model because of −0.99 rank correlation within the pairs (ANHPRM, ANHCOMP) and (HALPRM, HALCOMP).
[c] Standardized regression coefficients (SRCs) in final regression model.
[d] Cumulative R^2 value with entry of each variable into regression model.
[e] Standardized rank regression coefficients (SRRCs) in final regression model.

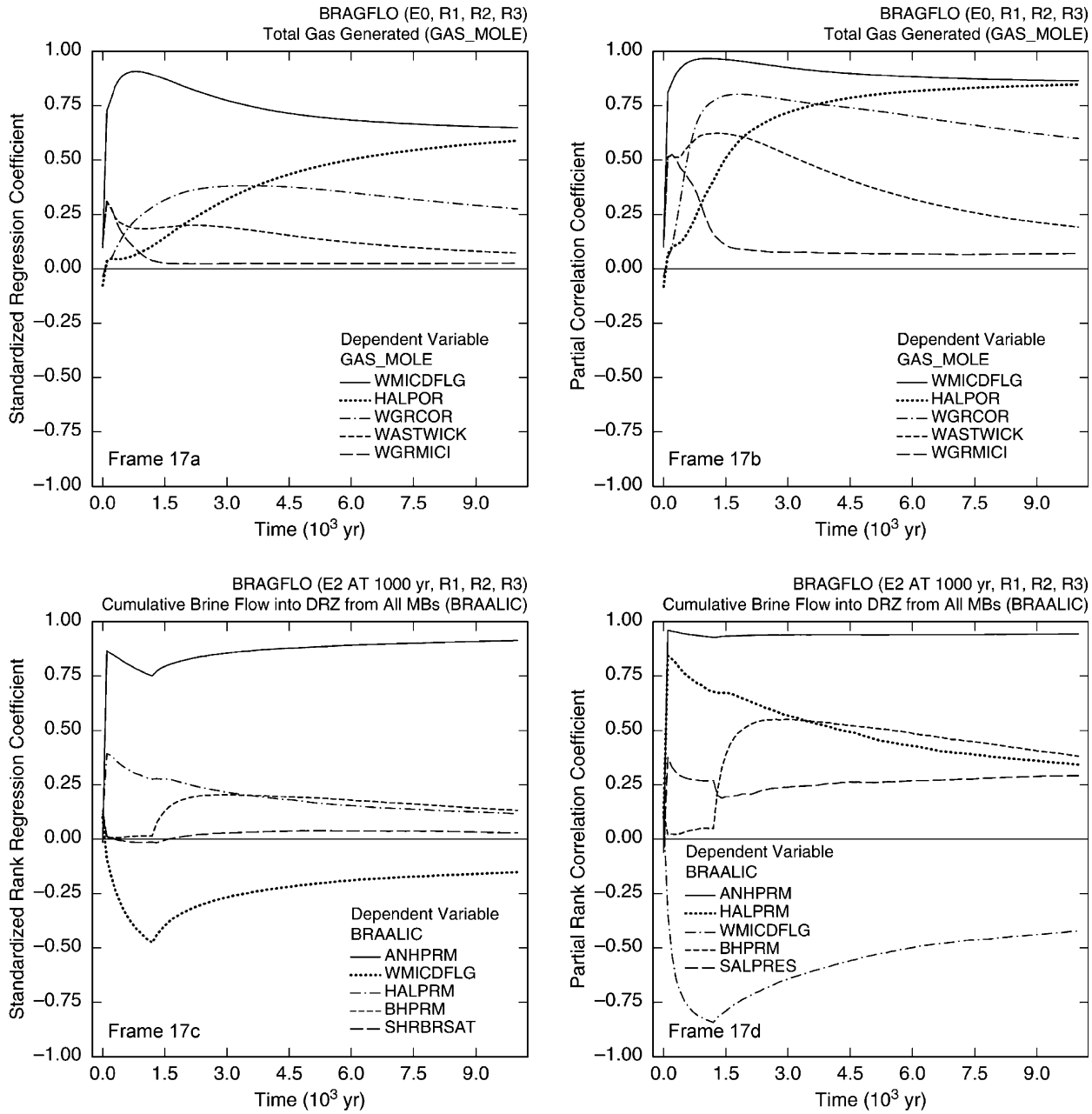

Fig. 17. Time-dependent coefficients: (a) and (b) SRCs and PCCs for cumulative gas generation under undisturbed (i.e. E0) conditions (y = E0:GAS_MOLE; see Fig. 11b); and (c) and (d), SRRC's and PRCCs for cumulative brine flow into DRZ with an E2 intrusion at 1000 yr (y = E2:BRAALIC; see Fig. 11c).

The F-statistic and the Kruskal–Wallis statistic are then used to indicate if the y values associated with these intervals appear to have different means and distributions, respectively. The χ^2-statistic for CMDs involves a further partitioning of the y values into values above and below the median for all y in Eq. (2.7) (i.e. the horizontal line in Fig. 18a), with the corresponding significance test used to indicate if the y values associated with the individual intervals defined for x_j appear to have medians that are different from the median for all values of y. The χ^2-statistic for SI involves a partitioning of the y values in Eq. (2.7) into intervals of equal probability analogous to the partitioning of the values of x_j (i.e. the horizontal lines in Fig. 18b), with the corresponding significance test used to indicate if the distribution of the points (x_{ij}, y_i) over the cells in Fig. 18b appears to be different from what would be expected if there was no relationship between x_j and y. For each statistic, a p-value can be calculated which corresponds to the probability of observing a stronger pattern than the one actually observed if there is no relationship between x_j and y. An ordering of p-values then provides a ranking of variable importance (i.e. the smaller the p-value, the stronger the effect of x_j on y appears to be). More detail on these and other related procedures is given in Kleijnen and Helton [255,256].

As an example, analyses for y = E2:PORVOL_T with the tests for CMNs, CMDs, CLs and SI are presented in Table 6.

Table 5
Stepwise regression analyses with raw and rank-transformed data with pooled results for replicates R1, R2 and R3 (i.e. for a total of 300 observations) for output variable E2:PORVOL_T at 10,000 yr

Step[a]	Raw data: y = E2:PORVOL_T			Rank-transformed data: y = E2:PORVOL_T		
	Variable[b]	SRC[c]	R^{2d}	Variable[b]	SRRC[e]	R^{2d}
1	HALPRM	0.37	0.15	HALPRM	0.35	0.13
2	BHPRM	0.33	0.25	ANHPRM	0.23	0.18
3	ANHPRM	0.24	0.31	HALPOR	0.13	0.20
4	HALPOR	0.15	0.33			

[a] Steps in stepwise regression analysis with significance levels of $\alpha = 0.02$ and $\alpha = 0.05$ required of a variable for entry into and retention in a regression model, respectively.
[b] Variables listed in order of selection in regression analysis with ANHCOMP and HALCOMP excluded from entry into regression model because of -0.99 rank correlation within the pairs (ANHPRM, ANHCOMP) and (HALPRM, HALCOMP).
[c] Standardized regression coefficients (SRCs) in final regression model.
[d] Cumulative R^2 value with entry of each variable into regression model.
[e] Standardized rank regression coefficients (SRRCs) in final regression model.

For perspective, tests based on p-values for CCs and RCCs are also presented in Table 6, with the p-values indicating the probability of observing larger, in absolute value, CCs and RCCs due to chance variation in the absence of any relationship between x_j and y [255]. The ordering of variable importance with CMNs, CMDs, CLs and SI is different from the orderings obtained with CCs and RCCs. In particular, the tests for CMNs, CMDs, CLs and SI are identifying the nonlinear and nonmonotonic relationship involving BHPRM that is being missed with the tests based on CCs and RCCs. If desired, the top-down correlation technique introduced by Iman and Conover could be used to provide a formal assessment of the agreement between the results for the different sensitivity analysis procedures in Table 6 [255,257].

Variance decomposition procedures provide another way to identify nonlinear and nonmonotonic relationships and are typically implemented with Monte Carlo procedures (Section 2.4). In addition, many procedures have been proposed by the ecological community for identifying nonrandom patterns that may have a use in sampling-based sensitivity analysis (e.g. Refs. [258–271]). Finally, the two-dimensional Kolmogorov–Smirnov test has the potential to be a useful tool for the identification of nonrandom patterns in sampling-based sensitivity analysis (e.g. Refs. [272–275]). Further information on sampling-based procedures for uncertainty and sensitivity analysis is available in a number of reviews (e.g. Refs. [32,38,217,255,276–282]).

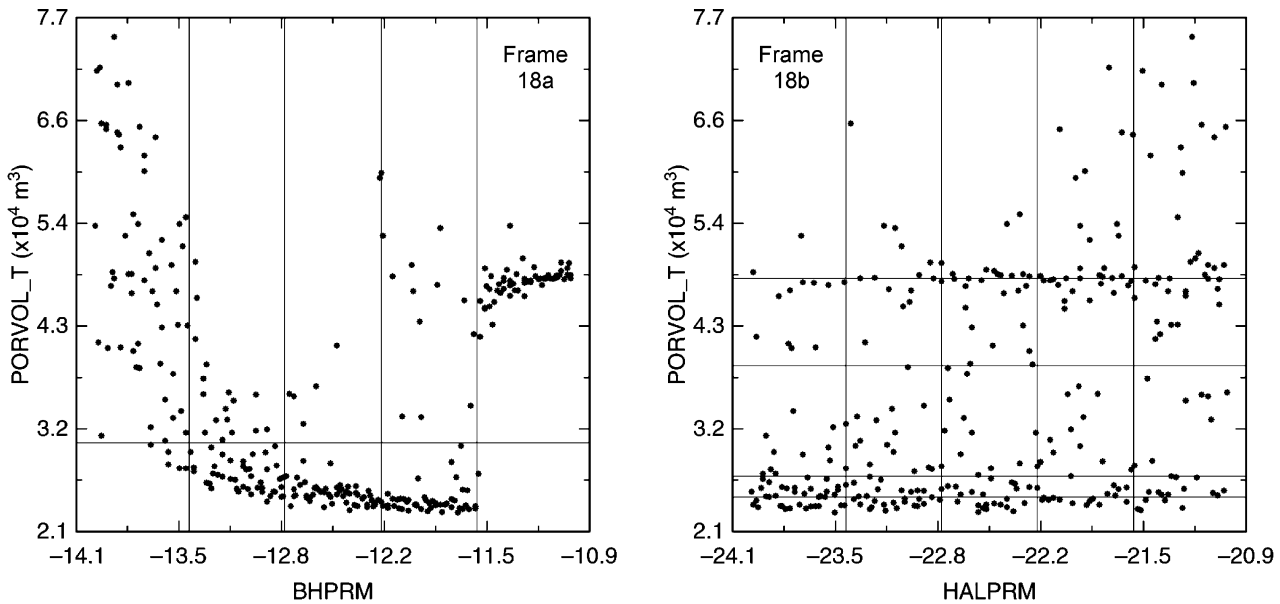

Fig. 18. Partitionings of (x_{ij}, y_i), $i = 1, 2, \ldots, nS = 300$: (a) division of x_j = BHPRM into intervals of equal probability and y = E2:PORVOL_T into values above and below the median, and (b) division of x_j = HALPRM and y = E2:PORVOL_T into intervals of equal probability.

Table 6
Sensitivity results based on CMNs, CMDs, CLs, SI, CCs and RCCs for y = E2:PORVOL_T

Variable	CMN		CMD		CL		SI		CC		RCC	
	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val	Rank	p-Val
BHPRM	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000	10.0	0.3295	4.0	0.0926
HALPRM	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0001	1.0	0.0000	1.0	0.0000
ANHPRM	3.0	0.0005	3.0	0.0007	3.0	0.0000	4.0	0.0082	2.0	0.0000	2.0	0.0000
HALPOR	4.0	0.0341	6.0	0.0700	5.0	0.1072	5.0	0.1137	3.0	0.0097	3.0	0.0225
ANHBCEXP	5.0	0.0496	5.0	0.0595	4.0	0.0655	18.5	0.5739	9.0	0.1938	9.0	0.2535
ANHBCVGP	6.0	0.0899	16.0	0.4884	6.0	0.1248	13.0	0.2942	4.0	0.0894	5.0	0.1248
SHRBRSAT	9.0	0.1923	7.0	0.0823	8.0	0.1464	7.0	0.1850	21.0	0.6859	15.0	0.4559
BPPRM	10.0	0.2010	4.0	0.0477	7.0	0.1350	18.5	0.5739	22.0	0.7069	14.0	0.4329
WGRCOR	19.0	0.5386	17.0	0.5249	10.0	0.2320	3.0	0.0003	14.0	0.4688	18.0	0.6601

7. Uncertainty in analyses for complex systems (adapted from Ref. [276], Chapt. 10)

7.1. Stochastic and subjective uncertainty

Many large analyses maintain a separation between two categorizations of uncertainty: (i) stochastic uncertainty, which arises because the system under study can behave in many different ways (e.g. many different accidents are possible at a nuclear power station), and (ii) subjective uncertainty, which arises from a lack of knowledge about quantities assumed to have fixed values in a particular analysis (e.g. a reactor containment building might be assumed to have a fixed failure strength, with the exact value of this strength being unknown). Thus, stochastic uncertainty in a property of the system under study, and subjective uncertainty is a property of the analysis and the associated analysts. Alternative terminology includes the use of aleatory, variability, irreducible and type A as alternatives to the designation stochastic and the use of epistemic, state of knowledge, reducible and type B as alternatives to the designation subjective. The categorization and treatment of stochastic and subjective uncertainty in analyses for complex systems has been widely discussed from a variety of perspectives [7,8,283–295]. Further, the use of probability to characterize both subjective and stochastic uncertainty can be traced back to the beginnings of the formal development of probability in the late seventeenth century [296–298].

The distributions in Eq. (1.2) were assumed to characterize subjective uncertainty, and the probability space associated with these distributions was represented by ($\mathscr{S}_{su}$, $\mathfrak{S}_{su}$, p_{su}), with the subscript su used as a designation for subjective. Analyses that involve stochastic and subjective uncertainty have two underlying probability spaces: a probability space ($\mathscr{S}_{st}$, $\mathfrak{S}_{st}$, p_{st}) for stochastic uncertainty, and a probability space ($\mathscr{S}_{su}$, $\mathfrak{S}_{su}$, p_{su}) for subjective uncertainty. In the preceding, the subscript 'st' is used as a designator for 'stochastic'.

An example of a large analysis that maintained a separation between stochastic and subjective uncertainty is the NRC's reassessment of the risk from commercial nuclear reactors in the United States (i.e. NUREG-1150), where stochastic uncertainty arose from the many possible accidents that could occur at the power plants under study and subjective uncertainty arose from the many uncertain quantities required in the estimation of the probabilities and consequences of these accidents [172,173,183]. Numerous other examples also exist (e.g. Refs. [184–187,299–308]).

7.2. Performance assessment for the WIPP

This presentation will use the PA carried out in support of the DOE's 1996 CCA for the WIPP as an example of an analysis involving both stochastic and subjective uncertainty [190,191,249]. Parts of this analysis involving the model for two-phase flow implemented in the BRAGFLO program have already been introduced and used to illustrate uncertainty and sensitivity analysis in the presence of subjective uncertainty (Section 6.1). Although the analyses with BRAGFLO were an important part of the 1996 WIPP PA, they constitute only one component of a large analysis. The following provides a high-level overview of sampling-based uncertainty and sensitivity analysis in the 1996 WIPP PA. The need to treat both stochastic and subjective uncertainty in the 1996 WIPP PA arose from regulations promulgated by the EPA and briefly summarized in the next paragraph.

The following is the central requirement in the EPA's regulation for the WIPP, 40 CFR 191, Subpart B, and the primary determinant of the conceptual and computational structure of the 1996 WIPP PA (p. 38086, Ref. [192]):

§ 191.13 Containment requirements:

(a) Disposal systems for spent nuclear fuel or high-level or transuranic radioactive wastes shall be designed to provide a reasonable expectation, based upon performance assessments, that cumulative releases of

radionuclides to the accessible environment for 10,000 years after disposal from all significant processes and events that may affect the disposal system shall: (1) Have a likelihood of less than one chance in 10 of exceeding the quantities calculated according to Table 1 (Appendix A);[1] and (2) Have a likelihood of less than one chance in 1,000 of exceeding ten times the quantities calculated according to Table 1 (Appendix A).

(b) Performance assessments need not provide complete assurance that the requirements of 191.13(a) will be met. Because of the long time period involved and the nature of the events and processes of interest, there will inevitably be substantial uncertainties in projecting disposal system performance. Proof of the future performance of a disposal system is not to be had in the ordinary sense of the word in situations that deal with much shorter time frames. Instead, what is required is a reasonable expectation, on the basis of the record before the implementing agency, that compliance with 191.13(a) will be achieved.The EPA also promulgated 40 CFR 194 [193], where the following elaboration on the intent of 40 CFR 191.13 is given (pp. 5242–5243, Ref. [193]):

§ 194.34 Results of performance assessments.

(a) The results of performance assessments shall be assembled into "complementary, cumulative distribution functions" (CCDFs) that represent the probability of exceeding various levels of cumulative release caused by all significant processes and events. (b) Probability distributions for uncertain disposal system parameter values used in performance assessments shall be developed and documented in any compliance application. (c) Computational techniques, which draw random samples from across the entire range of the probability distributions developed pursuant to paragraph (b) of this section, shall be used in generating CCDFs and shall be documented in any compliance application. (d) The number of CCDFs generated shall be large enough such that, at cumulative releases of 1 and 10, the maximum CCDF generated exceeds the 99th percentile of the population of CCDFs with at least a 0.95 probability. (e) Any compliance application shall display the full range of CCDFs generated. (f) Any compliance application shall provide information which demonstrates that there is at least a 95 percent level of statistical confidence that the mean of the population of CCDFs meets the containment requirements of § 191.13 of this chapter.

In addition to the requirements in 40 CFR 191.13 and 40 CFR 194.34 just quoted, 40 CFR 191 and 40 CFR 194 contain many additional requirements for the certification of the WIPP for the disposal of TRU waste [309]. However, it is the indicated requirements that determine the overall structure of the 1996 WIPP PA.

Together, 191.13(a) and 194.34(a) lead to a CCDF and boundary line [310–312] as illustrated in Fig. 19, with the CCDF for releases to the accessible environment required to fall below the boundary line. The CCDF derives from disruptions that could occur in the future and is thus characterizing the effects of stochastic uncertainty. In contrast, 194.34(b) and (c) require the characterization and propagation of the effects of subjective uncertainty. Ultimately, this uncertainty will lead to a distribution of CCDFs of the form illustrated in Fig. 19, with this distribution deriving from subjective uncertainty.

The probability space $(\mathscr{S}_{su}, \mathfrak{S}_{su}, p_{su})$ for subjective uncertainty used in the 1996 WIPP PA has already been introduced in Section 6.1, with Table 1 listing examples of the 57 uncertain variables associated with the elements $\mathbf{x}_{su}$ of $\mathscr{S}_{su}$. Specifically, $\mathbf{x}_{su}$ is a vector of the form

$$\mathbf{x}_{su} = [x_1, x_2, \ldots, x_{57}] \tag{7.1}$$

in the 1996 WIPP PA. The probability space $(\mathscr{S}_{su}, \mathfrak{S}_{su}, p_{su})$ was defined by specifying distributions for the elements of $\mathbf{x}_{su}$ as indicated in Eq. (1.2) and illustrated in Fig. 10.

In the 1996 WIPP PA, the probability space $(\mathscr{S}_{st}, \mathfrak{S}_{st}, p_{st})$ for stochastic uncertainty derives from the many different disruptions that could occur at the WIPP over the 10,000 yr regulatory time frame imposed on it. In particular, regulatory guidance [309] and extensive review of potential features, events and processes (FEPs) that could affect the WIPP [314] led to the elements $\mathbf{x}_{st}$ of the sample space $\mathscr{S}_{st}$

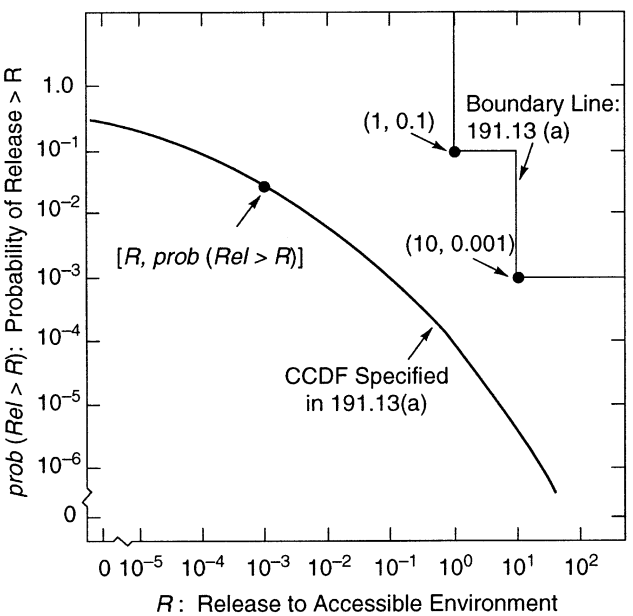

Fig. 19. Boundary line and associated CCDF specified in 40 CFR 191, Subpart B (Fig. 2, Ref. [313]).

[1] Radionuclide releases normalized to amount of radioactive material placed in the disposal facility; see Refs. [192,194] for a description of the normalization process.

being defined as vectors of the form

$$\mathbf{x}_{st} = [\underbrace{t_1, l_1, e_1, b_1, p_1, \mathbf{a}_1}_{\text{1st intrusion}}, \underbrace{t_2, l_2, e_2, b_2, p_2, \mathbf{a}_2}_{\text{2nd intrusion}}, \cdots, \underbrace{t_n, l_n, e_n, b_n, p_n, \mathbf{a}_n}_{n\text{th intrusion}}, t_{\min}], \tag{7.2}$$

where n is the number of exploratory drilling intrusions for natural resources (i.e. oil or gas) that occur in the immediate vicinity of the repository, t_i is the time (yr) of the ith intrusion, l_i designates the location of the ith intrusion, e_i designates the penetration of an excavated or nonexcavated area by the ith intrusion, b_i designates whether or not the ith intrusion penetrates pressurized brine in the Castile Formation, p_i designates the plugging procedure used with the ith intrusion (i.e. continuous plug, two discrete plugs, three discrete plugs), $\mathbf{a}_i$ designates the type of waste penetrated by the ith intrusion (i.e. no waste, contact-handled waste, and remotely handled waste, with $\mathbf{a}_i$ represented as a vector because a single drilling intrusion can penetrate several 'waste streams' that have different properties), and $t_{\min}$ is the time at which potash mining occurs within the land withdrawal boundary. The definition of $(\mathscr{S}_{st}, \mathfrak{S}_{st}, p_{st})$ was then completed by assigning a distribution to each element of $\mathbf{x}_{st}$ [315].

The FEPs review process also led to the identification of processes and associated models for use in the estimation of consequences (e.g. normalized radionuclide releases to the accessible environment in the context of the EPA regulations) for elements $\mathbf{x}_{st}$ of $\mathscr{S}_{st}$ (Fig. 20, Table 7). Symbolically, this estimation process can be represented by

$$\begin{aligned} f(\mathbf{x}_{st}) = {} & f_C(\mathbf{x}_{st}) + f_{SP}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})] \\ & + f_{DBR}\{\mathbf{x}_{st}, f_{SP}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})], f_B(\mathbf{x}_{st})\} + f_{MB}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})] \\ & + f_{DL}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})] + f_S[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})] \\ & + f_{S-T}\{\mathbf{x}_{st,0}, f_{S-F}(\mathbf{x}_{st,0}), f_{N-P}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]\}, \end{aligned} \tag{7.3}$$

where $f(\mathbf{x}_{st})$ ~ normalized radionuclide release to the accessible environment associated with $\mathbf{x}_{st}$ and, in general, many additional consequences, $\mathbf{x}_{st}$ ~ particular future under consideration, $\mathbf{x}_{st,0}$ ~ future involving no drilling intrusions but a mining event at the same time $t_{\min}$ as in $\mathbf{x}_{st}$, $f_C(\mathbf{x}_{st})$ ~ cuttings and cavings release to accessible environment for $\mathbf{x}_{st}$ calculated with CUTTINGS_S, $f_B(\mathbf{x}_{st})$ ~ results calculated for $\mathbf{x}_{st}$ with BRAGFLO (in practice, $f_B(\mathbf{x}_{st})$ is a vector containing a large amount of information including time-dependent pressures and saturations for gas and brine), $f_{SP}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]$ ~ spallings release to accessible environment for $\mathbf{x}_{st}$ calculated with the spallings model contained in CUTTINGS_S, $f_{DBR}\{\mathbf{x}_{st}, f_{SP}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})], f_B(\mathbf{x}_{st})\}$ ~ direct brine release to accessible environment for $\mathbf{x}_{st}$ calculated with a modified version of BRAGFLO designated BRAGFLO_DBR, $f_{MB}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]$ ~ release through anhydrite marker beds to accessible environment for $\mathbf{x}_{st}$ calculated with NUTS, $f_{DL}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]$ ~ release through Dewey Lake Red Beds to accessible environment for $\mathbf{x}_{st}$ calculated with NUTS, $f_S[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]$ ~ release to land surface due to brine flow up a plugged borehole for $\mathbf{x}_{st}$ calculated with NUTS or PANEL as appropriate, $f_{S-F}(\mathbf{x}_{st,0})$ ~ flow field calculated for $\mathbf{x}_{st,0}$ with SECOFL2D, $f_{N-P}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]$ ~ release to Culebra for $\mathbf{x}_{st}$ calculated with NUTS or PANEL as appropriate,

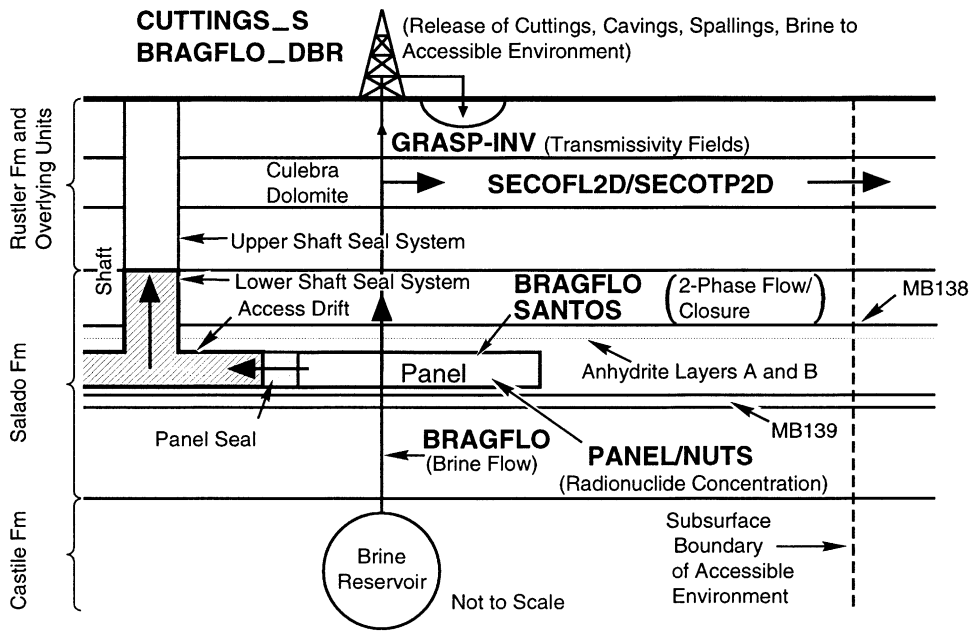

Fig. 20. Computer programs (models) used in 1996 WIPP PA (Fig. 5, Ref. [313]).

Table 7
Summary of computer models used in the 1996 WIPP PA (Table 1, Ref. [313])

BRAGFLO: Calculates multiphase flow of gas and brine through a porous, heterogeneous reservoir. Uses finite difference procedures to solve system of nonlinear partial differential equations that describes the mass conservation of gas and brine along with appropriate constraint equations, initial conditions and boundary conditions. Additional information: Section 4.2, Ref. [245]; Ref. [246]

BRAGFLO_DBR: Special configuration of BRAGFLO model used in calculation of dissolved radionuclide releases to the surface (i.e. direct brine releases) at the time of a drilling intrusion. Uses initial value conditions obtained from calculations performed with BRAGFLO and CUTTINGS_S. Additional information: Section 4.7, Ref. [245]; Ref. [316]

CUTTINGS_S: Calculates the quantity of radioactive material brought to the surface in cuttings and cavings and also in spallings generated by an exploratory borehole that penetrates a waste panel, where cuttings designates material removed by the drillbit, cavings designates material eroded into the borehole due to shear stresses resulting from the circular flow of the drilling fluid (i.e. mud), and spallings designates material carried to the borehole at the time of an intrusion due to the flow of gas from the repository to the borehole. Spallings calculation uses initial value conditions obtained from calculations performed with BRAGFLO. Additional information: Sects. 4.5, 4.6, Ref. [245]; Ref. [317]

GRASP-INV: Generates transmissivity fields (estimates of transmissivity values) conditioned on measured transmissivity values and calibrated to steady-state and transient pressure data at well locations using an adjoint sensitivity and pilot-point technique. Additional information: Refs. [318, 319]

NUTS: Solves system of partial differential equations for radionuclide transport in vicinity of repository. Uses brine volumes and flows calculated by BRAGFLO as input. Additional information: Section 4.3, Ref. [245]; Ref. [320]

PANEL: Calculates rate of discharge and cumulative discharge of radionuclides from a waste panel through an intruding borehole. Discharge is a function of fluid flow rate, elemental solubility and radionuclide inventory. Uses brine volumes and flows calculated by BRAGFLO as input. Based on solution of system of linear ordinary differential equations. Additional information: Section 4.4, Ref. [245]; Ref. [320]

SANTOS: Solves quasistatic, large deformation, inelastic response of two-dimensional solids with finite element techniques. Used to determine porosity of waste as a function of time and cumulative gas generation, which is an input to calculations performed with BRAGFLO. Additional information: Section 4.2.3, Ref. [245]; Refs. [321,322]

SECOFL2D: Calculates single-phase Darcy flow for groundwater flow in two dimensions. The formulation is based on a single partial differential equation for hydraulic head using fully implicit time differencing. Uses transmissivity fields generated by GRASP-INV. Additional information: Section 4.8, Ref. [245]; Ref. [323]

SECOTP2D: Simulates transport of radionuclides in fractured porous media. Solves two partial differential equations: one provides two-dimensional representation for convective and diffusive radionuclide transport in fractures and the other provides one-dimensional representation for diffusion of radionuclides into rock matrix surrounding the fractures. Uses flow fields calculated by SECOFL2D. Additional information: Section 4.9, Ref. [245]; Ref. [323]

and $f_{S-T}\{\mathbf{x}_{st,0}, f_{S-F}(\mathbf{x}_{st,0}), f_{N-P}[\mathbf{x}_{st}, f_B(\mathbf{x}_{st})]\}$ ~ groundwater transport release through Culebra to accessible environment calculated with SECOTP2D ($\mathbf{x}_{st,0}$ is used as an argument to f_{S-T} because drilling intrusions are assumed to cause no perturbations to the flow field in the Culebra).

The probability space $(\mathscr{S}_{st}, \mathscr{F}_{st}, p_{st})$ for stochastic uncertainty and the function f indicated in Eq. (7.3) lead to the required CCDF for normalized releases to the accessible environment (Fig. 19). In particular, this CCDF can be represented as an integral involving $(\mathscr{S}_{st}, \mathscr{F}_{st}, p_{st})$ and f (Fig. 21). If $(\mathscr{S}_{st}, \mathscr{F}_{st}, p_{st})$ and f could be unambiguously defined, then the CCDF in Fig. 21 could be determined with certainty and compared against the specified boundary line. Unfortunately, such certainty does not exist in the 1996 WIPP PA, which leads to the probability space $(\mathscr{S}_{su}, \mathscr{F}_{su}, p_{su})$ for subjective uncertainty.

When the elements $\mathbf{x}_{su}$ of $\mathscr{S}_{su}$ are included, the function f in Eq. (7.3) has the form $f(\mathbf{x}_{st}, \mathbf{x}_{su})$. In turn, the expression defining the CCDF in Fig. 21 becomes

$$\text{prob}(Rel > R|\mathbf{x}_{su}) = \int_{\mathscr{S}_{st}} \delta_R[f(\mathbf{x}_{st}, \mathbf{x}_{su})] d_{st}(\mathbf{x}_{st}|\mathbf{x}_{su}) dV_{st}, \quad (7.4)$$

where $\delta_R[f(\mathbf{x}_{st}, \mathbf{x}_{su})] = 1$ if $f(\mathbf{x}_{st}, \mathbf{x}_{su}) > R$ and 0 if $f(\mathbf{x}_{st}, \mathbf{x}_{su}) \le R$. Uncertainty in $\mathbf{x}_{su}$ as characterized by $(\mathscr{S}_{su}, \mathscr{F}_{su}, p_{su})$ then leads to a distribution of CCDFs, with one CCDF resulting for each $\mathbf{x}_{su}$ in $\mathscr{S}_{su}$ (Fig. 22).

7.3. Implementation of 1996 WIPP PA

The guidance in 194.34(a) was implemented by developing the probability space $(\mathscr{S}_{st}, \mathscr{F}_{st}, p_{st})$, the function $f(\mathbf{x}_{st}, \mathbf{x}_{su})$, and a Monte Carlo procedure based on simple random sampling (Section 5.1) for the approximation of the integral, and hence the associated CCDF, in Eq. (7.4). Conditional on an element $\mathbf{x}_{su}$ of $\mathscr{S}_{su}$, the Monte Carlo

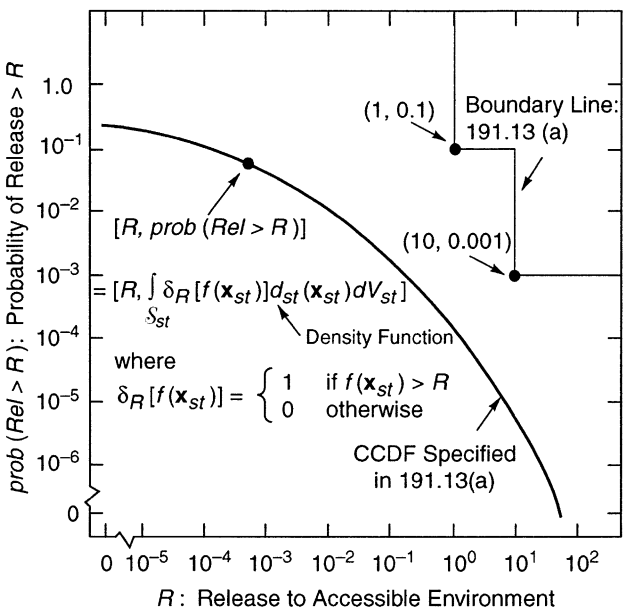

Fig. 21. Definition of CCDF specified in 40 CFR 191, Subpart B as an integral involving the probability space $(\mathscr{S}_{st}, \mathscr{F}_{st}, p_{st})$ for stochastic uncertainty and a function f defined on $\mathscr{S}_{st}$ (Fig. 4, Ref. [313]).

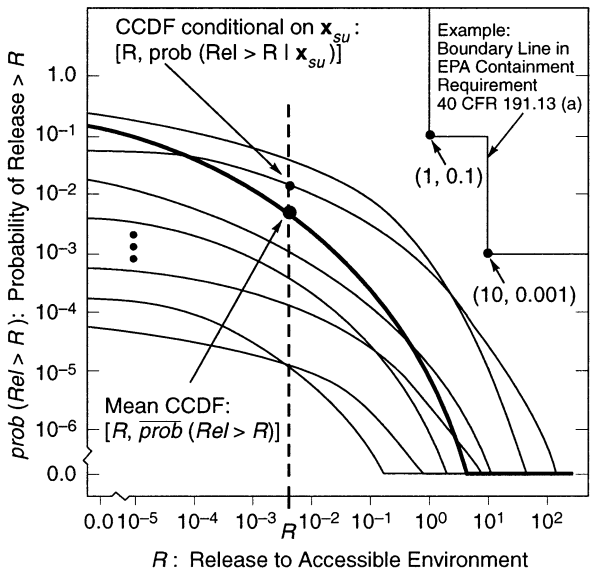

Fig. 22. Individual CCDFs conditional on elements $\mathbf{x}_{su}$ of $\mathscr{S}_{su}$ (i.e. CCDFs represented by $[R, \text{prob}(Rel > R|\mathbf{x}_{su})]$; see Eq. (7.4)) and associated mean CCDF (i.e. CCDF represented by $[R, \overline{\text{prob}}(Rel > R)]$; see Eq. (7.7)).

approximation procedure has the form

$$\text{prob}(Rel > R|\mathbf{x}_{su}) \doteq \sum_{i=1}^{nS} \delta_R[f(\mathbf{x}_{st,i}, \mathbf{x}_{su})]/nS, \tag{7.5}$$

where $\mathbf{x}_{st,i}$, $i = 1, 2, \ldots, nS = 10,000$, is a random sample from $(\mathscr{S}_{st}, \mathfrak{S}_{st}, p_{st})$. This approximation procedure required evaluating the models in Table 7 for a relatively small number of elements of $\mathscr{S}_{st}$ and then using these evaluations to the construct $f(\mathbf{x}_{st,i}, \mathbf{x}_{su})$ for the large number of sample elements (i.e. $nS = 10,000$) used in the summation in Eq. (7.5) (see Refs. [245,315–317,320,323,324] for numerical details).

The guidance in 194.34(b) was implemented by developing the probability space $(\mathscr{S}_{su}, \mathfrak{S}_{su}, p_{su})$. Latin hypercube sampling was selected as the sampling technique required in 194.34(c) because of the efficient manner in which it stratifies across the range of each sampled variable. For a Latin hypercube or random sample of size n, the requirement in 194.34(c) is equivalent to the inequality

$$1 - 0.99^n > 0.95, \tag{7.6}$$

which results in a minimum value of 298 for n. In consistency with the preceding result, the 1996 WIPP PA used an LHS of size 300 from the probability space $(\mathscr{S}_{su}, \mathfrak{S}_{su}, p_{su})$ for subjective uncertainty. Actually, as discussed below, three replicated LHSs of size 100 each were used, which resulted in a total sample size of 300 (Section 6.1). Further, the requirement in 194.34(d) is met by simply providing plots that contain all the individual CCDFs produced in the analysis (i.e. one CCDF for each LHS element, which generates plots of the form indicated in Fig. 22).

The requirement in 194.34(f) involves the mean of the distribution of CCDFs, with this distribution resulting from subjective uncertainty (Fig. 22). In particular, each individual CCDF in Fig. 22 is conditional on an element $\mathbf{x}_{su}$ of $\mathscr{S}_{su}$ and is defined by the points $[R, \text{prob}(Rel > R|\mathbf{x}_{su})]$, with $\text{prob}(Rel > R|\mathbf{x}_{su})$ given in Eq. (7.5). Similarly, the mean CCDF is defined by the points $[R, \overline{\text{prob}}(Rel > R)]$, where

$$\begin{aligned}
&\overline{\text{prob}}(Rel > R) \\
&\quad = \text{mean probability of a release greater than size } R \\
&\quad = \int_{\mathscr{S}_{su}} \text{prob}(Rel > R|\mathbf{x}_{su}) d_{su}(\mathbf{x}_{su}) \text{d}V_{su} \\
&\quad = \int_{\mathscr{S}_{su}} \left\{ \int_{\mathscr{S}_{st}} \delta_R[f(\mathbf{x}_{st}, \mathbf{x}_{su})] d_{st}(\mathbf{x}_{st}|\mathbf{x}_{su}) \text{d}V_{st} \right\} d_{su}(\mathbf{x}_{su}) \text{d}V_{su}
\end{aligned} \tag{7.7}$$

and $d_{su}(\mathbf{x}_{su})$ is the density function associated with $(\mathscr{S}_{su}, \mathfrak{S}_{su}, p_{su})$. The integral over $\mathscr{S}_{su}$ in the definition of $\overline{\text{prob}}(Rel > R)$ is too complex to be determined exactly. The EPA anticipated that a sampling-based integration procedure would be used to estimate this integral, with the requirement in 194.34(f) placing a condition on the accuracy of this procedure.

Given that Latin hypercube sampling is to be used to estimate the outer integral in Eq. (7.7), the confidence intervals required in 194.34(f) can be obtained with the replicated sampling technique proposed by Iman (Section 5.3). As discussed in Section 5.3, the LHS to be used is repeatedly generated with different random seeds. These samples lead to a sequence $\overline{\text{prob}}_r(Rel > R)$, $r = 1, 2, \ldots, nR$, of estimated mean exceedance probabilities, where $\overline{\text{prob}}_r(Rel > R)$ defines the mean CCDF obtained for sample r (i.e. $\overline{\text{prob}}_r(Rel > R)$ is the mean probability that a normalized release of size R will be exceeded; see Eq. (7.7)) and nR is the number of independent LHSs generated with different random seeds. Then,

$$\overline{\overline{\text{prob}}}(Rel > R) = \sum_{r=1}^{nR} \overline{\text{prob}}_r(Rel > R)/nR \tag{7.8}$$

and

$$\text{SE}(R) = \left\{ \sum_{r=1}^{nR} \left[\overline{\overline{\text{prob}}}(Rel > R) - \overline{\text{prob}}_r(Rel > R) \right]^2 / nR(nR-1) \right\}^{1/2} \tag{7.9}$$

provide an additional estimate of the mean CCDF and estimates of the standard errors associated with the individual mean exceedance probabilities $\overline{\overline{\text{prob}}}(Rel > R)$ that define this CCDF. The t-distribution with $nR-1$ degrees of freedom can be used to place confidence intervals around the mean exceedance probabilities for individual R values (i.e. around $\overline{\overline{\text{prob}}}(Rel > R)$). Specifically, the $1-\alpha$ confidence interval is given by $\overline{\overline{\text{prob}}}(Rel > R) \pm t_{1-\alpha/2}\text{SE}(R)$, where $t_{1-\alpha/2}$ is the $1-\alpha/2$ quantile of

the t-distribution with $nR-1$ degrees of freedom (e.g. $t_{1-\alpha/2}=4.303$ for $\alpha=0.05$ and $nR=3$; Ref. [218], Table A25). The same procedure can also be used to place pointwise confidence intervals around percentile curves. The implementation of this procedure is the reason for the three replicated LHSs indicated in Section 6.1.

At the beginning of the computational implementation of the 1996 WIPP PA, only the 31 variables in $\mathbf{x}_{su}$ that are used as input to BRAGFLO had been fully specified (i.e. their distributions D_j had been unambiguously defined); the remaining variables that would be incorporated into the definition of $\mathbf{x}_{su}$ were still under development. To allow the calculations with BRAGFLO to proceed, the LHSs indicated in Section 6.1 were actually generated from $nX = 75$ variables, with the first 31 variables being the then specified inputs to BRAGFLO and the remaining 44 variables being assigned uniform distributions on [0,1]. Later, when the additional variables were fully specified, the uniformly distributed variables were used to generate sampled values from them consistent with their assigned distributions. This procedure allowed the analysis to go forward while maintaining the integrity of the Latin hypercube sampling procedure for the overall analysis. As previously indicated, 26 additional variables were eventually defined, with the result that the elements $\mathbf{x}_{su}$ of $\mathscr{S}_{su}$ had an effective dimension of $nX = 57$.

7.4. Uncertainty and sensitivity analysis results in 1996 WIPP PA

The CCDF used in comparisons with the EPA release limits (Figs. 19 and 21) is the most important single result generated in the 1996 WIPP PA. This CCDF arises from stochastic uncertainty. However, because there is subjective uncertainty in quantities used in the generation of this CCDF, its value cannot be unambiguously known. The use of Latin hypercube sampling leads to an estimate of the uncertainty in the location of this CCDF (Fig. 23), with the individual CCDFs falling substantially to the left of the release limits. The left frame (Fig. 23a) shows the individual CCDFs obtained for replicate R1, and the right frame (Fig. 23b) shows the mean and selected percentile curves obtained from pooling the three replicates. The mean curve in Fig. 23b is formally defined in Eq. (7.7), and the construction procedures used to obtain the individual curves in Fig. 23b are described in conjunction with Fig. 13.

The replicated samples described in Section 6.1 were used to obtain an indication of the stability of results obtained with Latin hypercube sampling. For the total release CCDFs in Fig. 23, the results obtained for the three replicates (i.e. R1, R2, R3) were very stable, with little variation in the locations of the mean and percentile curves occurring across replicates (Fig. 24a). Indeed, the mean and percentile curves for the individual replicates overlie each other to the extent that they are almost indistinguishable. As a result, the procedure indicated in conjunction with Eqs. (7.8) and (7.9) provides a very tight confidence interval around the estimated mean CCDF (Fig. 24b).

The sampling-based approach to uncertainty analysis has created a pairing between the individual LHS elements and the individual CCDFs in Fig. 23a that can be explored with the previously discussed sensitivity analysis techniques (Section 6). One possibility for

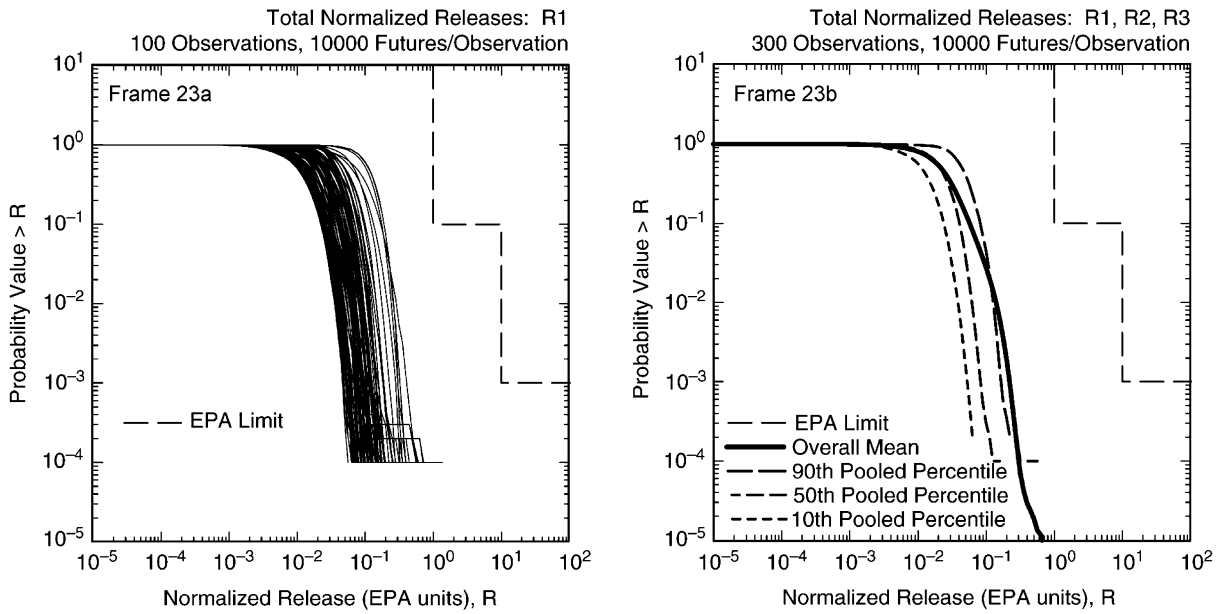

Fig. 23. Distribution of CCDFs for total normalized release to the accessible environment over 10,000 yr: (a) 100 individual CCDFs for replicate R1, and (b) mean and percentile curves estimated from 300 CCDFs obtained by pooling replicates R1, R2 and R3 (Figs. 6 and 7, Ref. [313]).

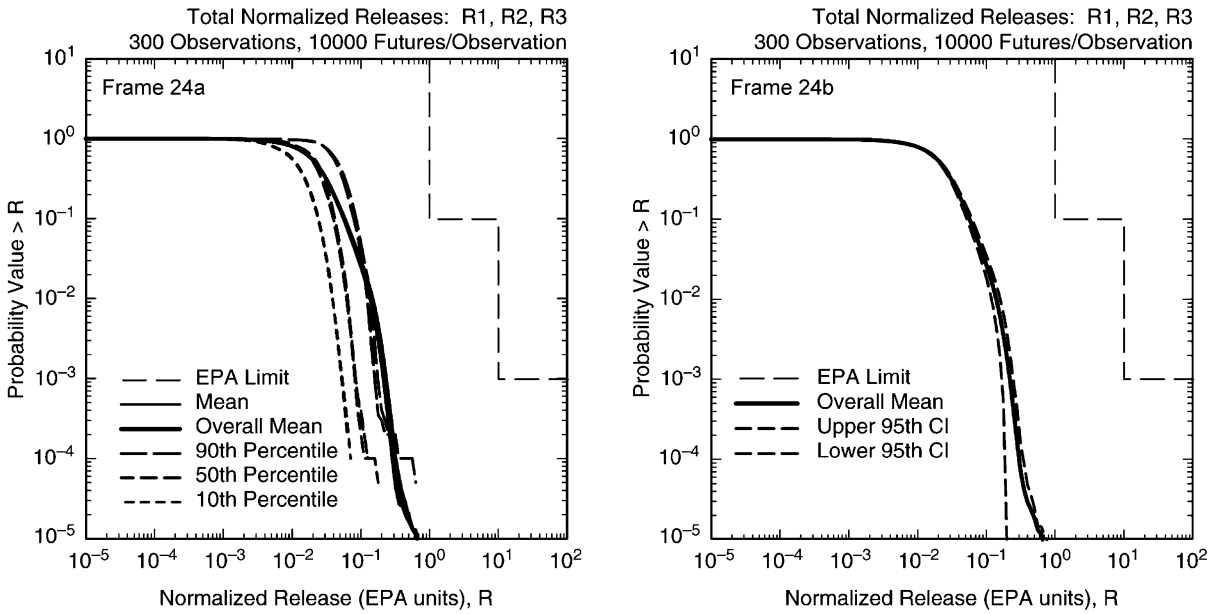

Fig. 24. Stability of estimated distribution of CCDFs for normalized release to the accessible environment: (a) mean and percentile curves for individual replicates, and (b) confidence interval around mean CCDF obtained by pooling the three individual replicates (Fig. 8, Ref. [313]).

investigating the sources of the uncertainty that give rise to the distribution of CCDFs in Fig. 23a is to determine what is giving rise to the variation in exceedance probabilities for individual release values on the abscissa. This variation in exceedance probabilities can be investigated in exactly the same manner as the variation in cumulative gas generation (GAS_MOLE) and brine inflow (BRAALIC) at individual times was investigated for the curves in Fig. 11 and presented in Fig. 17. Specifically, PRCCs, SRRCs, or some other measure of sensitivity can be calculated for the exceedance probabilities associated with individual release values. This measure for different sampled variables can be plotted above the corresponding release values on the abscissa and then connected to obtain a representation for how sensitivity changes for changing values on the abscissa. For the CCDFs in Fig. 23a, this analysis approach shows that the exceedance probabilities for individual release values are primarily influenced by WMICDFLG and WTAUFAIL, with the exceedance probabilities tending to increase as WMICDFLG increases and tending to decrease as WTAUFAIL increases (Fig. 25).

Another possibility is to reduce the individual CCDFs to expected values over stochastic uncertainty and then to perform a sensitivity analysis on the resultant expected values. In the context of the CCDF representation in Eq. (7.4), this expected value can be formally defined by

$$E(R|\mathbf{x}_{su}) = \int_{\mathscr{S}_{st}} f(\mathbf{x}_{st}, \mathbf{x}_{su}) d_{st}(\mathbf{x}_{st}|\mathbf{x}_{su}) \mathrm{d}V_{st}. \quad (7.10)$$

The LHS then results in a sequence of values $E(R|\mathbf{x}_{su,k})$, $k = 1, 2, \ldots, n\mathrm{LHS} = 300$, that can be explored with the previously discussed sensitivity analysis procedures. For example, stepwise regression analysis shows that WMICDFLG and WTAUFAIL are the dominant variables with respect to the uncertainty in $E(R|\mathbf{x}_{su})$, with lesser effects due to a number of additional variables (Table 8).

This section briefly describes the 1996 WIPP PA and illustrates uncertainty and sensitivity analysis procedures based on Latin hypercube sampling in the context of this PA. Additional details are available in other presentations [191,245,249,325].

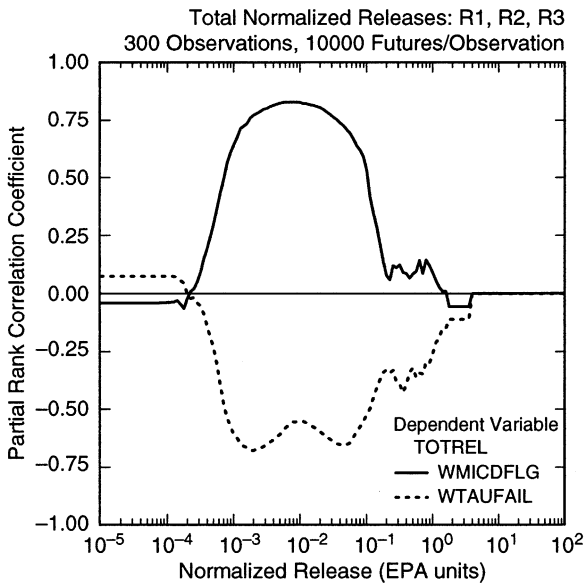

Fig. 25. Sensitivity analysis based on PRCCs for CCDFs for normalized release to the accessible environment (Fig. 14, Ref. [325]).

Table 8
Stepwise regression analysis with rank-transformed data for expected normalized release associated with individual CCDFs for total release due to cuttings and cavings, spallings and direct brine release (Table 5, Ref. [325])

Step[a]	Expected normalized release		
	Variable[b]	SRRC[c]	R^{2d}
1	WMICDFLG	0.60	0.40
2	WTAUFAIL	−0.39	0.55
3	WGRCOR	0.21	0.59
4	WPRTDIAM	−0.19	0.63
5	HALPOR	0.17	0.65
6	BHPRM	−0.17	0.68
7	HALPRM	0.16	0.71
8	WASTWICK	0.11	0.72
9	ANHPRM	0.09	0.73

[a] Steps in stepwise regression analysis with significance levels of $\alpha = 0.02$ and $\alpha = 0.05$ required of a variable for entry into and retention in a regression model, respectively.
[b] Variables listed in order of selection in regression analysis with ANHCOMP and HALCOMP excluded from entry into regression model because of −0.99 rank correlation within the pairs (ANHPRM, ANHCOMP) and (HALPRM, HALCOMP).
[c] Standardized rank regression coefficients (SRRCs) in final regression model.
[d] Cumulative R^2 value with entry of each variable into regression model.

8. Discussion

Latin hypercube sampling has become a widely used sampling technique for the propagation of uncertainty in analyses of complex systems. A check of the original article [31] in Science Citation Index or SciSearch can be used to obtain both a list of all citations and also the most recent citations to this technique. This review ends with a discussion of some of the reasons for the popularity of Latin hypercube sampling (Section 8.1) and some additional thoughts on the propagation of uncertainty in analyses for complex systems (Section 8.2).

8.1. Popularity of Latin hypercube sampling

Reasons that have led to the popularity of Monte Carlo techniques in general and Latin hypercube sampling in particular for uncertainty and sensitivity analysis of complex models include (i) conceptual simplicity and ease of implementation, (ii) dense stratification over the range of each sampled variable, (iii) direct provision of uncertainty analysis results without the use of surrogate models as approximations to the original model, (iv) availability of a variety of sensitivity analysis procedures, and (v) effectiveness as a model verification procedure. The preceding reasons are discussed in more detail below.

Conceptual simplicity and ease of implementation. A Monte Carlo approach to the propagation of uncertainty is easy to explain. Further, the definition of Latin hypercube sampling is straightforward, and the reason why its enforced stratification improves the results of an analysis for a given sample size is easy to grasp on an intuitive level. Thus, the presentation of Monte Carlo and Latin hypercube results to individuals of different levels of technical sophistication (e.g. other scientists working in the same or related fields, private or governmental decision makers, the general public) is relatively straightforward. In contrast, some of the other techniques for the propagation and analysis of uncertainty are less transparent (e.g. RSM, FAST, Sobol' variance decomposition, FPI) and thus more difficult to present.

Analyses based on Latin hypercube sampling are typically easy to implement. Software is available to generate LHSs and also to implement the Iman/Conover restricted pairing technique for the control of correlations within the sample (e.g. Ref. [165]). Further, propagation of the sample through the model under consideration is straightforward in most analyses. In practice, this propagation often involves little more than putting a 'DO Loop' around the model which (i) reads the individual sample elements, (ii) uses these elements to generate input in the form required by the model, (iii) runs the model, and (iv) saves model results for later analysis.

In contrast, implementation of the other analysis procedures can be considerably more difficult: (i) RSM requires the development of both a suitable experimental design and the construction of a surrogate model, (ii) differential analysis requires the determination of the necessary model derivatives, (iii) FAST and Sobol' variance decomposition require the development and evaluation of suitable integrals involving the model to obtain the associated variance decompositions, and (iv) FPI requires the evaluation and use of model derivatives in the location of the MPP. Not only are the above procedures conceptually and computationally complex but, in many analyses, they can require more computational effort (i.e. model evaluations) than a Monte Carlo analysis with Latin hypercube sampling.

Analyses that involve a single model are relatively easy to implement and explain. Analyses that involve a sequence of linked, and possibly quite complex, models are more difficult to implement and explain. Examples of such analyses are the NRC's reassessment of the risk from commercial nuclear power reactors (i.e. NUREG-1150) [172,173,183] and the DOE's PA in support of a CCA for the WIPP [191,245,249,325]. However, in such analyses, a sampling-based approach provides a way to examine results at model interfaces and develop a computational strategy for the overall assembly of the analysis. Analyses using the other techniques described in Section 2 seem less useful in the design, integration and ultimate performance of an analysis that involves the propagation of uncertainty through a sequence of linked models.

Dense stratification over range of each sampled variable. Latin hypercube sampling results in a denser stratification over the range of each sampled variable than would be obtained with a classical experimental design of the type

typically used in conjunction with RSM and a more uniform stratification than would be obtained with random sampling. Further, the random pairing associated with Latin hypercube sampling spreads the sampled points throughout the high-dimensional sample space.

Real analyses typically have a large number of analysis outcomes of interest. Further, these outcomes are often spatially or temporally dependent. The result is that most, if not all, of the sampled variables can be important to one or more of the analysis outcomes. The dense stratification over the range of each sampled variable with Latin hypercube sampling results in each variable being sampled in a manner that allows its effects to be recognized if such effects exist.

It is a mistake to assume that the important effects associated with a variable only occur at the end points of its range. Instead, it is quite possible that the most important effects associated with a variable could occur in an interior part of its range (e.g. Fig. 18a). The dense stratification associated with Latin hypercube sampling allows the identification of such effects when they occur. Further, this stratification also facilitates the identification of interactions involving multiple variables (e.g. Fig. 16; also Figs. 8 and 9, Ref. [171]).

Direct provision of uncertainty analysis results. Because probabilistic weights can be associated with individual sample elements, Latin hypercube sampling, random sampling and stratified sampling can be used to obtain estimates of distribution functions directly from model results. Further, these estimates are unbiased, although some bias may be introduced if the Iman/Conover restricted pairing technique (Section 5.1) is used.

Latin hypercube sampling tends to produce more stable results (i.e. less variation in estimated distribution functions from sample to sample) than random sampling. However, examples can be obtained in which Latin hypercube sampling and random sampling produce results of similar stability by constructing a model in which variations in model behavior take place on a scale that is much smaller than the interval sizes in the LHS that result from the sample size selected for use. Stratified sampling can produce better distribution function estimates than either Latin hypercube or random sampling provided enough information is available to define the strata and calculate the associate strata probabilities. Thus, stratified sampling is typically used only when a substantial knowledge base has already been obtained about the problem under consideration and is usually not appropriate in an initial exploratory analysis. Further, it is difficult to define a meaningful stratified sampling plan when many analysis outcomes are under consideration, as is usually the case in most real analyses.

In contrast to Latin hypercube, random and stratified sampling, FPI is intended primarily for estimating the tails of a distribution rather than the full distribution. Differential analysis in conjunction with the associated Taylor series provides an estimate for model variance rather than the full distribution function; further, the expected values of analysis outcomes are usually taken to be the outcome of the model evaluated at the expected values of the inputs. The FAST approach and Sobol' variance decomposition are also used to estimate expected values and variances rather than full distribution functions, although the calculations used to obtain expected values can also be used to produce estimated distribution functions.

An important characteristic of Latin hypercube and random sampling is that the resultant model evaluations can be used to provide estimated distribution functions for all analysis outcomes. In particular, a different analysis/computational strategy does not have to be developed and implemented for each analysis outcome. As already indicated, real analyses typically have a large number of outcomes of interest, and the necessity to develop a separate investigation for each of them can impose unreasonable demands on both human and computational resources.

Variety of sensitivity analysis procedures. Latin hypercube and random sampling generate a mapping from uncertain analysis inputs to analysis results. Once generated, this mapping can be explored with a variety of techniques, including examination of scatterplots, correlation analysis, regression analysis, rank transformations, tests for nonmonotonic patterns, and tests for random patterns. This variety of techniques allows flexibility in developing a sensitivity analysis that is appropriate for the particular analysis situation under consideration. Again, Latin hypercube sampling is particularly effective in sensitivity analyses with small samples due to its efficient stratification across the range of each uncertain variable.

Sensitivity analyses in differential analysis and RSM are typically based on assessing either the effects of perturbations away from base values or fractional contributions to variance. In either case, the resultant sensitivity analyses are no better than the surrogate models (i.e. Taylor series or response surfaces) over the range of uncertainty under consideration. Fast probability integration (FPI) is primarily an uncertainty analysis procedure and is usually not used in sensitivity analysis.

The FAST approach and Sobol' variance decomposition provide very appealing sensitivity analysis results. In particular, they provide a complete decomposition of variance into the components due to individual variables and interactions between variables. Unfortunately, if the model under consideration is expensive to evaluate or a large number of analysis outcomes are being investigated, the computational cost of implementing these procedures may be prohibitive. Although the FAST approach and the Sobol' variance decomposition are calculated under the assumption that model inputs are independent, variance decomposition procedures exist that can be used with correlated inputs [135–137].

Model verification. Sampling-based uncertainty and sensitivity analysis provides a very powerful tool in model verification. Here, model verification is used to mean checking the correctness of the implementation of a model

and/or an analysis and thus is distinct from model validation, which involves checking the capability of the model and/or analysis to represent the physical system under study. Propagation of a sample through an analysis provides a very extensive exercising of its many components. Gross errors will often be revealed by the failure of the analysis for some sample elements or possibly by the appearance of clearly erroneous results. Further, subtler errors are often revealed in sensitivity analyses (e.g. a variable having a small negative effect when the underlying physics implies that it should have a small positive effect or when a variable is shown to affect an analysis result on which it should have no effect).

Sensitivity analysis provides a way to examine a large number of analysis outcomes for anomalous behavior in a very rapid manner. Further, relatively small effects can be observed. Sampling-based sensitivity analysis is much more effective in searching for analysis errors than simply running the model for a limited number of cases and then examining the results of these calculations. A sampling-based sensitivity analysis should be included as part of any serious model/analysis verification effort. Latin hypercube sampling is particularly effective in model verification due to the dense stratification across the range of each sampled variable.

8.2. Additional thoughts

Uncertainty and sensitivity analyses for complex systems are typically iterative. An initial study is often performed to gain perspectives on (i) the behavior of the model(s) involved, (ii) strategies for carrying out a final and defensible analysis, and (iii) the most important variables with respect to the uncertainty in outcomes of interest. In such preliminary analyses, rather crude characterizations of variable uncertainty may be adequate. Once system behavior is better understood and the more important variables with respect to this behavior are identified, resources can be focused on improving the characterization of the uncertainty in these important variables. Further, iterative analyses facilitate quality assurance by providing repeated opportunities to check the correctness of model and analysis implementation. Sampling-based approaches to uncertainty and sensitivity analysis are particularly effective in iterative analyses due to the extensive exercising of the model(s) and associated analysis structure and the availability of a variety of uncertainty and sensitivity analysis results.

Concern is often expressed about the computational cost of carrying out a Monte Carlo analysis. In most analyses, the human cost of developing the model, carrying out the analysis, and documenting and defending the analysis will be far greater than the computational cost (i.e. for CPU time) of performing the necessary calculations. Latin hypercube sampling was developed to improve the quality of uncertainty and sensitivity analysis results relative to those that could be obtained with a random sample of the same size. However, if the model is inexpensive to evaluate, a large sample size can be used, and whether Latin hypercube or random sampling is used will have little effect on either the cost of the analysis or the quality of the results obtained.

Some individuals express a broad, almost philosophical, dislike for Monte Carlo analysis procedures. This makes little sense. Monte Carlo procedures are just a way of carrying out a numerical integration and developing a mapping between model inputs and outputs. There may be reasons to question the model in use or the distributions assigned to uncertain variables, but these are not reasons to question Monte Carlo procedures themselves. Of course, a Monte Carlo analysis has to be carried out with a sufficiently large sample to produce results with a resolution appropriate for the purposes of the analysis. Replicated sampling provides one way to investigate the robustness of analysis outcomes and thus the appropriateness of the sample size selected for use; further, the individual replicates can be pooled to produce the final presentation results of the analysis.

Many large analyses involve a separation of subjective (i.e. epistemic) uncertainty and stochastic (i.e. aleatory) uncertainty (e.g. the NUREG-1150 probabilistic risk assessments (PRAs) and the WIPP PA as previously mentioned). In such analyses, a common strategy is to use Latin hypercube sampling to propagate the effects of subjective uncertainty, and random or stratified sampling to propagate the effects of stochastic uncertainty. With this approach, the effect of stochastic uncertainty is being calculated conditional on individual LHS elements. Typical analysis outcomes are distributions of CCDFs, with the individual CCDFs arising from stochastic uncertainty and the distributions of CCDFs arising from subjective uncertainty. The efficient stratification associated with Latin hypercube sampling is important in analyses of this type due to the possibly large computational effort required in the determination of the effects of stochastic uncertainty.

Random or stratified sampling is often a better choice than Latin hypercube sampling for the incorporation of stochastic uncertainty into an analysis. With random sampling, it is possible to build up a sample by selecting one sampled value at a time. In contrast, Latin hypercube sampling requires the entire sample to be selected at one time. As a result, random sampling often works better than Latin hypercube sampling when the values to be sampled are closely linked to effects that derive from previously sampled values (e.g. when the stopping point for a sampling process is determined by previously sampled values). The WIPP PA used random sampling to incorporate the effects of stochastic uncertainty due to the need to determine the effects of randomly occurring drilling intrusions over a 10,000 yr period. With stratified sampling, it is possible to force the inclusion of low-probability but high-consequence events into the analysis. The NUREG-1150 PRAs used stratified sampling implemented by event trees to assure the inclusion of, and also to calculate the probability of, low-probability/high-consequence accidents.

No approach to the propagation and analysis of uncertainty can be optimum for all needs. For example and depending on the particular problem, stratified sampling or FPI can be more appropriate than Latin hypercube sampling for the estimation of the extreme quantiles of a distribution. Likewise, differential analysis may be the preferred approach if it is necessary to determine the effects of small perturbations away from base-case values, and the FAST approach or Sobol' variance decomposition may be the preferred approach if it is necessary to determine a complete variance decomposition. However, it is the authors' view that Monte Carlo analysis with Latin hypercube sampling is the most broadly applicable approach to the propagation and analysis of uncertainty and often the only approach that is needed.

Acknowledgements

Work performed for Sandia National Laboratories (SNL), which is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000. Review provided at SNL by M. Chavez, S. Halliday, and J.W. Garner. Editorial support provided by F. Puffer, J. Ripple, and H. Radke of Tech Reps, Inc.

References

[1] NCRP (National Council on Radiation Protection and Measurements). A guide for uncertainty analysis in dose and risk assessments related to environmental contamination. NCRP Commentary No. 14, Bethesda, MD: National Council on Radiation Protection and Measurements; 1996.
[2] NRC (National Research Council). Science and judgment in risk assessment. Washington, DC: National Academy Press; 1994.
[3] NRC (National Research Council). Issues in risk assessment. Washington, DC: National Academy Press; 1993.
[4] US EPA (US Environmental Protection Agency). An SAB report: multi-media risk assessment for radon, Review of uncertainty analysis of risks associated with exposure to radon. EPA-SAB-RAC-93-014, Washington, DC: US Environmental Protection Agency; 1993.
[5] IAEA (International Atomic Energy Agency). Evaluating the reliability of predictions made using environmental transfer models. Safety Series No. 100, Vienna: International Atomic Energy Agency; 1989.
[6] Beck MB. Water-quality modeling: a review of the analysis of uncertainty. Water Resour Res 1987;23(8):1393–442.
[7] Helton JC, Burmaster DE. Guest editorial: treatment of aleatory and epistemic uncertainty in performance assessments for complex systems. Reliab Engng Syst Saf 1996;54(2–3):91–4.
[8] Helton JC. Uncertainty and sensitivity analysis in the presence of stochastic and subjective uncertainty. J Stat Comput Simul 1997; 57(1–4):3–76.
[9] Cook I, Unwin SD. Controlling principles for prior probability assignments in nuclear risk assessment. Nuclear Sci Engng 1986; 94(2):107–19.
[10] Mosleh A, Bier VM, Apostolakis G. A critique of current practice for the use of expert opinions in probabilistic risk assessment. Reliab Engng Syst Saf 1988;20(1):63–85.
[11] Hora SC, Iman RL. Expert opinion in risk analysis: the NUREG-1150 methodology. Nuclear Sci Engng 1989;102(4):323–31.
[12] Svenson O. On expert judgments in safety analyses in the process industries. Reliab Engng Syst Saf 1989;25(3):219–56.
[13] Keeney RL, von Winterfeldt D. Eliciting probabilities from experts in complex technical problems. IEEE Trans Engng Manage 1991; 38(3):191–201.
[14] Bonano EJ, Hora SC, Keeney RL, von Winterfeldt D. Elicitation and use of expert judgment in performance assessment for high-level radioactive waste repositories. Albuquerque: Sandia National Laboratories; 1990.
[15] Bonano EJ, Apostolakis GE. Theoretical foundations and practical issues for using expert judgments in uncertainty analysis of high-level radioactive waste disposal. Radioact Waste Manage Nuclear Fuel Cycle 1991;16(2):137–59.
[16] Cooke RM. Experts in uncertainty: opinion and subjective probability in science. Oxford: Oxford University Press; 1991.
[17] Meyer MA, Booker JM. Eliciting and analyzing expert judgment: a practical guide. New York: Academic Press; 1991.
[18] Ortiz NR, Wheeler TA, Breeding RJ, Hora SC, Myer MA, Keeney RL. Use of expert judgment in NUREG-1150. Nuclear Engng Des 1991;126(3):313–31.
[19] Chhibber S, Apostolakis G, Okrent D. A taxonomy of issues related to the use of expert judgments in probabilistic safety studies. Reliab Engng Syst Saf 1992;38(1–2):27–45.
[20] Kaplan S. Expert information versus expert opinions: another approach to the problem of eliciting combining using expert knowledge in PRA. Reliab Engng Syst Saf 1992;35(1):61–72.
[21] Otway H, von Winterfeldt D. Expert judgement in risk analysis and management: process, context, and pitfalls. Risk Anal 1992;12(1): 83–93.
[22] Thorne MC, Williams MMR. A review of expert judgement techniques with reference to nuclear safety. Prog Nuclear Saf 1992;27(2–3):83–254.
[23] Thorne MC. The use of expert opinion in formulating conceptual models of underground disposal systems and the treatment of associated bias. Reliab Engng Syst Saf 1993;42(2–3):161–80.
[24] Evans JS, Gray GM, Sielken Jr. RL, Smith AE, Valdez-Flores C, Graham JD. Use of probabilistic expert judgement in uncertainty analysis of carcinogenic potency, part 1. Regul Toxicol Pharmacol 1994;20(1):15–36.
[25] Budnitz RJ, Apostolakis G, Boore DM, Cluff LS, Coppersmith KJ, Cornell CA, Morris PA. Use of technical expert panels: applications to probabilistic seismic hazard analysis. Risk Anal 1998;18(4): 463–9.
[26] Goossens LHJ, Harper FT. Joint EC/USNRC expert judgement driven radiological protection uncertainty analysis. J Radiol Prot 1998;18(4):249–64.
[27] Siu NO, Kelly DL. Bayesian parameter estimation in probabilistic risk assessment. Reliab Engng Syst Saf 1998;62(1–2):89–116.
[28] Goossens LHJ, Harper FT, Kraan BCP, Metivier H. Expert judgement for a probabilistic accident consequence uncertainty analysis. Radiat Prot Dosimetry 2000;90(3):295–301.
[29] McKay M, Meyer M. Critique of and limitations on the use of expert judgements in accident consequence uncertainty analysis. Radiat Prot Dosimetry 2000;90(3):325–30.
[30] Feller W. An introduction to probability theory and its applications, vol. 2. 2nd ed. New York: Wiley; 1971.
[31] McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 1979;21(2):239–45.
[32] Iman RL. Uncertainty and sensitivity analysis for computer modeling applications. In: Cruse TA, editor. Reliability technology—1992. The Winter Annual Meeting of the American Society of

Mechanical Engineers, Anaheim, California, November, 8–13, 1992 vol. 28. New York, NY: American Society of Mechanical Engineers, Aerospace Division; 1992. p. 153–68.
[33] MacDonald RC, Campbell JE. Valuation of supplemental and enhanced oil recovery projects with risk analysis. J Petrol Technol 1986;38(1):57–69.
[34] Breshears DD, Kirchner TB, Whicker FW. Contaminant transport through agroecosystems: assessing relative importance of environmental, physiological, and management factors. Ecol Appl 1992; 2(3):285–97.
[35] Ma JZ, Ackerman E. Parameter sensitivity of a model of viral epidemics simulated with Monte Carlo techniques. II. Durations and peaks. Int J Biomed Comput 1993;32(3–4):255–68.
[36] Ma JZ, Ackerman E, Yang J-J. Parameter sensitivity of a model of viral epidemics simulated with Monte Carlo techniques. I. Illness attack rates. Int J Biomed Comput 1993;32(3–4):237–53.
[37] Whiting WB, Tong T-M, Reed ME. Effect of uncertainties in thermodynamic data and model parameters on calculated process performance. Ind Engng Chem Res 1993;32(7):1367–71.
[38] Blower SM, Dowlatabadi H. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. Int Stat Rev 1994;62(2):229–43.
[39] Gwo JP, Toran LE, Morris MD, Wilson GV. Subsurface stormflow modeling with sensitivity analysis using a Latin-Hypercube sampling technique. Ground Water 1996;34(5):811–8.
[40] Helton JC, Anderson DR, Baker BL, Bean JE, Berglund JW, Beyeler W, Economy K, Garner JW, Hora SC, Iuzzolino HJ, Knupp P, Marietta MG, Rath J, Rechard RP, Roache PJ, Rudeen DK, Salari K, Schreiber JD, Swift PN, Tierney MS, Vaughn P. Uncertainty and sensitivity analysis results obtained in the 1992 performance assessment for the waste isolation pilot plant. Reliab Engng Syst Saf 1996;51(1):53–100.
[41] Chan MS. The consequences of uncertainty for the prediction of the effects of schistosomiasis control programmes. Epidemiol Infect 1996;11(3):537–50.
[42] Sanchez MA, Blower SM. Uncertainty and sensitivity analysis of the basic reproductive rate: tuberculosis as an example. Am J Epidemiol 1997;145(12):1127–37.
[43] Blower SM, Gershengorn HB, Grant RM. A tale of two futures: HIV and antiretroviral therapy in San Francisco. Science 2000;287(5453): 650–4.
[44] Cohen C, Artois M, Pontier D. A discrete-event computer model of feline Herpes virus within cat populations. Preventative Vet Med 2000;45(3–4):163–81.
[45] Kolev NI, Hofer E. Uncertainty and sensitivity analysis of a postexperiment simulation of nonexplosive melt-water interaction. Exp Ther Fluid Sci 1996;13(2):98–116.
[46] Caswell H, Brault S, Read AJ, Smith TD. Harbor porpoise and fisheries: an uncertainty analysis of incidental mortality. Ecol Appl 1998;8(4):1226–38.
[47] Hofer E. Sensitivity analysis in the context of uncertainty analysis for computationally intensive models. Comput Phys Commun 1999; 117(1–2):21–34.
[48] Metropolis N, Ulam S. The Monte Carlo method. J Am Stat Assoc 1949;44(247):335–41.
[49] Hammersley JM, Handscomb DC. Monte Carlo methods. London: Methuen; 1964.
[50] Teichroew D. A history of distribution sampling prior to the era of the computer and its relevance to simulation. J Am Stat Assoc 1965; 60:27–49.
[51] Halton JH. A retrospective and prospective survey of the Monte Carlo method. SIAM Rev 1970;12(1):1–63.
[52] James F. Monte Carlo theory and practice. Rep Prog Phys 1980; 43(9):1145–89.
[53] Rubinstein RY. Simulation and the Monte Carlo method. New York: Wiley; 1981.
[54] Turner JE, Wright HA, Hamm RN. A Monte Carlo primer for health physicists. Health Phys 1985;48(6):717–33.
[55] Kalos MH, Whitlock PA. Monte Carlo methods. New York: Wiley; 1986.
[56] Rief H, Gelbard EM, Schaefer RW, Smith KS. Review of Monte Carlo techniques for analyzing reactor perturbations. Nuclear Sci Engng 1986;92(2):289–97.
[57] Fishman GS. Monte Carlo: concepts, algorithms, and applications. New York: Springer; 1996.
[58] Albano EV. The Monte Carlo simulation method: a powerful tool for the study of reaction processes. Heterogeneous Chem Rev 1996;3(4): 389–418.
[59] Binder K. Applications of Monte Carlo methods to statistical physics. Rep Prog Phys 1997;60(5):487–559.
[60] Hurtado JE, Barbat AH. Monte Carlo techniques in computational stochastic mechanics. Arch Comput Meth Engng 1998;5(1):3–29.
[61] Hahn GJ, Shapiro SS. Statistical models in engineering. New York: Wiley; 1967.
[62] Tukey JW. The propagation of errors, fluctuations and tolerances: basic generalized formulas. Report No. 10, Princeton, NJ: Statistical Techniques Research Group, Princeton University Press; 1957.
[63] Tukey JW. The propagation of errors, fluctuations and tolerances: supplementary formulas. Report No. 11, Princeton, NJ: Statistical Techniques Research Group, Princeton University Press; 1957.
[64] Tukey JW. The propagation of errors, fluctuations and tolerances: an exercise in partial differentiation. Report No. 12, Princeton, NJ: Statistical Techniques Research Group, Princeton University Press; 1957.
[65] Tomovic R, Vukobratovic M. General sensitivity theory. New York: Elsevier; 1972.
[66] Frank PM. Introduction to system sensitivity theory. New York: Academic Press; 1978.
[67] Lewins J, Becker M, editors. Sensitivity and uncertainty analysis of reactor performance parameters. Advances in nuclear science and technology, vol. 14. New York: Plenum Press; 1982.
[68] Rabitz H, Kramer M, Dacol D. Sensitivity analysis in chemical kinetics. In: Rabinovitch BS, Schurr JM, Strauss HL, editors. Annual review of physical chemistry, vol. 34. Palo Alto, CA: Annual Reviews Inc; 1983. p. 419–61.
[69] Ronen Y. Uncertainty analysis. Boca Raton, FL: CRC Press; 1988.
[70] Turányi T. Sensitivity analysis of complex kinetic systems. Tools and applications. J Math Chem 1990;5(3):203–48.
[71] Cacuci DG, Weber CF, Oblow EM, Marable JH. Sensitivity theory for general systems of nonlinear equations. Nuclear Sci Engng 1980; 75(1):88–110.
[72] Cacuci DG. Sensitivity theory for nonlinear systems. I. Nonlinear functional analysis approach. J Math Phys 1981;22(12):2794–802.
[73] Cacuci DG. Sensitivity theory for nonlinear systems. II. Extensions to additional classes of responses. J Math Phys 1981;22(12): 2803–12.
[74] Cacuci DG, Schlesinger ME. On the application of the adjoint method of sensitivity analysis to problems in the atmospheric sciences. Atmósfera 1994;7(1):47–59.
[75] Dougherty EP, Rabitz H. A computational algorithm for the Green's function method of sensitivity analysis in chemical kinetics. Int J Chem Kinet 1979;11(12):1237–48.
[76] Dougherty EP, Hwang JT, Rabitz H. Further developments and applications of the Green's function method of sensitivity analysis in chemical kinetics. J Chem Phys 1979;71(4):1794–808.
[77] Hwang J-T, Dougherty EP, Rabitz S, Rabitz H. The Green's function method of sensitivity analysis in chemical kinetics. J Chem Phys 1978;69(11):5180–91.
[78] Vuilleumier L, Harley RA, Brown NJ. First- and second-order sensitivity analysis of a photochemically reactive system (a Green's function approach). Environ Sci Technol 1997;31(4):1206–17.
[79] Griewank A, Corliss G, editors. Automatic differentiation of algorithms: theory, implementation, and application. Proceedings

of the First SIAM Workshop on Automatic Differentiation, Breckenridge, Colorado, Philadelphia: Society for Industrial and Applied Mathematics; 1991.
[80] Berz M, Bischof C, Corliss G, Griewank A. Computational differentiation: techniques, applications, and tools. Philadelphia: Society for Industrial and Applied Mathematics; 1996.
[81] Bischof C, Khademi P, Mauer A, Carle A. Adifor 2.0: automatic differentiation of Fortran 77 programs. IEEE Comput Sci Engng 1996;3(3):18–32.
[82] Griewank A, Juedes D, Utke J. Algorithm 755: ADOL-C: a package for the automatic differentiation of algorithms written in C/C++. ACM Trans Math Software 1996;22(2):131–67.
[83] Bischof CH, Roh L, Mauer-Oats AJ. ADIC: an extensible automatic differentiation tool for ANSI-C. Software: Pract Experience 1997; 27(12):1427–56.
[84] Carmichael GR, Sandu A, Potra FA. Sensitivity analysis for atmospheric chemistry models via automatic differentiations. Atmos Environ 1997;31(3):475–89.
[85] Giering R, Kaminski T. Recipes for adjoint code construction. ACM Trans Math Software 1998;24(4):437–74.
[86] Tolsma JE, Barton PI. On computational differentiation. Comput Chem Engng 1998;22(4–5):475–90.
[87] Eberhard P, Bischof C. Automatic differentiation of numerical integration algorithms. Math Comput 1999;68(226):717–31.
[88] Griewank A. Evaluating derivatives: principles and techniques of algorithmic differentiation. Philadelphia: Society for Applied and Industrial Mathematics; 2000.
[89] Andres TH. Sampling methods and sensitivity analysis for large parameter sets. J Stat Comput Simul 1997;57(1–4):77–110.
[90] Kleijnen JPC. Sensitivity analysis and related analyses: a review of some statistical techniques. J Stat Comput Simul 1997;57(1–4): 111–42.
[91] Bates RA, Buck RJ, Riccomagno E, Wynn HP. Experimental design and observation for large systems. J R Stat Soc Ser B—Methodol 1996;58(1):77–94.
[92] Draper NR, Pukeisheim F. An overview of design of experiments. Stat Pap 1996;37(1):1–32.
[93] Draper NR, Lin DKJ. Response surface designs. Handbook of statistics, vol. 13. New York: Elsevier; 1996. pp. 343–75.
[94] Koehler JR, Owen AB. Computer experiments. Handbook of statistics, vol. 13. New York: Elsevier; 1996. pp. 261–308.
[95] Morris MD, Mitchell TJ. Exploratory designs for computational experiments. J Stat Plann Inference 1995;43(3):381–402.
[96] Bowman KP, Sacks J, Chang Y-F. Design and analysis of numerical experiments. J Atmos Sci 1993;50(9):1267–78.
[97] Kleijnen JPC. Sensitivity analysis of simulation experiments: regression-analysis and statistical design. Math Comput Simul 1992;34(3–4):297–315.
[98] Morris MD. Factorial sampling plans for preliminary computational experiments. Technometrics 1991;33(2):161–74.
[99] Currin C, Mitchell T, Morris M, Ylviskaker D. Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments. J Am Stat Assoc 1991;86(416): 953–63.
[100] Sacks J, Welch WJ, Mitchel TJ, Wynn HP. Design and analysis of computer experiments. Stat Sci 1989;4(4):409–35.
[101] Sacks J, Schiller SB, Welch WJ. Designs for computer experiments. Technometrics 1989;31(1):41–7.
[102] Cryer SA, Havens PL. Regional sensitivity analysis using a fractional factorial method for the USDA model GLEAMS. Environ Modell Software 1999;14(6):613–24.
[103] Rao GP, Sarkar PK. Sensitivity studies of air scattered neutron dose from particle accelerators. J Stat Comput Simul 1997;57(1–4): 261–70.
[104] Kleijnen JPC, van Ham G, Rotmans J. Techniques for sensitivity analysis of simulation models: a case study of the CO2 Greenhouse effect. Simulation 1992;58(6):410–7.
[105] Engel RE, Sorensen JM, May RS, Doran KJ, Trikouros NG, Mozias ES. Response surface development using RETRAN. Nuclear Technol 1991;93(1):65–81.
[106] Aceil SM, Edwards DR. Sensitivity analysis of thermal-hydraulic parameters and probability estimation of boiling transition in a standard BWR/6. Nuclear Technol 1991;93(2):123–9.
[107] Lee SH, Kim JS, Chang SH. A study on uncertainty and sensitivity of operational and modelling parameters for feedwater line break analysis. J Kor Nuclear Soc 1987;19(1):10–21.
[108] Kim HK, Lee YW, Kim TW, Chang SH. A procedure for statistical thermal margin analysis using response surface method and Monte Carlo technique. J Kor Nuclear Soc 1986;18(1):38–47.
[109] Myers RH. Response surface methodology—current status and future directions. J Qual Technol 1999;31(1):30–44.
[110] Myers RH, Khuri AI, Carter J, Walter H. Response surface methodology: 1966–1988. Technometrics 1989;31(2):137–57.
[111] Morton RH. Response surface methodology. Math Sci 1983;8: 31–52.
[112] Mead R, Pike DJ. A review of response surface methodology from a biometric viewpoint. Biometrics 1975;31:803–51.
[113] Hill WJ, Hunter WG. A review of response surface methodology: a literature review. Technometrics 1966;8(4):571–90.
[114] Myers RH. Response surface methodology. Boston, MA: Allyn and Bacon; 1971.
[115] Box GEP, Draper NR. Empirical model-building and response surfaces. New York: Wiley; 1987.
[116] Khuri AI, Cornell JA. Response surfaces: designs and analyses. New York: Marcel Dekker; 1987.
[117] Kleijnen JPC. Statistical tools for simulation practitioners. New York: Marcel Dekker; 1987.
[118] Cukier RI, Fortuin CM, Shuler KE, Petschek AG, Schaibly JH. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients, I. Theory. J Chem Phys 1973;59(8):3873–8.
[119] Schaibly JH, Shuler KE. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients, II. Applications. J Chem Phys 1973;59(8):3879–88.
[120] Cukier RI, Levine HB, Shuler KE. Nonlinear sensitivity analysis of multiparameter model systems. J Comput Phys 1978;26(1): 1–42.
[121] Sobol' IM. Sensitivity estimates for nonlinear mathematical models. Math Model Comput Exp 1993;1(4):407–14.
[122] Chan K, Saltelli A, Tarantola S. Winding stairs: a sampling tool to compute sensitivity indices. Stat Comput 2000;10(3):187–96.
[123] Jansen MJW. Analysis of variance designs for model output. Comput Phys Commun 1999;117(1–2):35–43.
[124] Jansen MJW, Rossing WAH, Daamen RA. Monte Carlo estimation of uncertainty contributions from several independent multivariate sources. In: Grasman J, van Straten G, editors. Predictability and nonlinear modeling in natural sciences and economics. Boston: Kluwer; 1994. p. 334–43.
[125] Rabitz H, Alis OF, Shorter J, Shim K. Efficient input–output model representations. Comput Phys Commun 1999;117(1–2):11–20.
[126] Saltelli A, Tarantola S, Chan KP-S. A quantitative model-independent method for global sensitivity analysis of model output. Technometrics 1999;41(1):39–56.
[127] Homma T, Saltelli A. Importance measures in global sensitivity analysis of nonlinear models. Reliab Engng Syst Saf 1996;52(1): 1–17.
[128] McRae GJ, Tilden JW, Seinfeld JH. Global sensitivity analysis—a computational implementation of the Fourier amplitude sensitivity test (FAST). Comput Chem Engng 1981;6(1):15–25.
[129] Saltelli A, Sobol' IM. About the use of rank transformation in sensitivity analysis of model output. Reliab Engng Syst Saf 1995; 50(3):225–39.
[130] Archer GEB, Saltelli A, Sobol' IM. Sensitivity measures, ANOVA-like techniques and the use of bootstrap. J Stat Comput Simul 1997; 58(2):99–120.

[131] Saltelli A, Bolado R. An alternative way to compute Fourier amplitude sensitivity test (FAST). Comput Stat Data Anal 1998; 26(4):267–79.
[132] Rabitz H, Alis OF. General foundations of high-dimensional model representations. J Math Chem 1999;25(2–3):197–233.
[133] Saltelli A, Tarantola S, Campolongo F. Sensitivity analysis as an ingredient of modeling. Stat Sci 2000;15(4):377–95.
[134] Cox DC. An analytic method for uncertainty analysis of nonlinear output functions, with applications to fault-tree analysis. IEEE Trans Reliab 1982;3(5):465–8.
[135] McKay MD. Evaluating prediction uncertainty. LA-12915-MS; NUREG/CR-6311, Los Alamos, NM: Los Alamos National Laboratory; 1995.
[136] McKay MD. Nonparametric variance-based methods of assessing uncertainty importance. Reliab Engng Syst Saf 1997;57(3):267–79.
[137] McKay MD, Morrison JD, Upton SC. Evaluating prediction uncertainty in simulation models. Comput Phys Commun 1999; 117(1–2):44–51.
[138] Haskin FE, Staple BD, Ding C. Efficient uncertainty analyses using fast probability integration. Nuclear Engng Des 1996;166(2): 225–48.
[139] Wu Y-T. Demonstration of a new, fast probability integration method for reliability analysis. J Engng Ind, Trans ASME, Ser B 1987;109(1):24–8.
[140] Schanz RW, Salhotra A. Evaluation of the Rackwitz-Fiessler uncertainty analysis method for environmental fate and transport method. Water Resour Res 1992;28(4):1071–9.
[141] Wu Y-T, Millwater HR, Cruse TA. Advanced probabilistic structural method for implicit performance functions. AIAA J 1990;28(9): 1663–9.
[142] Wu Y-T, Wirsching PH. New algorithm for structural reliability. J Engng Mech 1987;113(9):1319–36.
[143] Chen X, Lind NC. Fast probability integration by three-parameter normal tail approximation. Struct Saf 1983;1(4):169–76.
[144] Rackwitz R, Fiessler B. Structural reliability under combined random load sequences. Comput Struct 1978;9(5):489–94.
[145] Hasofer AM, Lind NC. Exact and invariant second-moment code format. J Engng Mech Div, Proc Am Soc Civil Eng 1974;100(EM1): 111–21.
[146] Castillo E, Solares C, Gómez P. Tail uncertainty analysis in complex systems. Artif Intell 1997;96(2):395–419.
[147] Castillo E, Sarabia JM, Solares C, Gómez P. Uncertainty analyses in fault trees and Bayesian networks using FORM/SORM methods. Reliab Engng Syst Saf 1999;65(1):29–40.
[148] Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical recipes in FORTRAN: the art of scientific computing, 2nd ed. Cambridge: Cambridge University Press; 1992.
[149] Barry TM. Recommendations on the testing and use of pseudo-random number generators used in Monte Carlo analysis for risk assessment. Risk Anal 1996;16(1):93–105.
[150] L'Ecuyer P. Random number generation. In: Banks J, editor. Handbook of simulation: principles, methodology, advances, applications, and practice. New York: Wiley; 1998. p. 93–137.
[151] Iman R, Conover WJ. A distribution-free approach to inducing rank correlation among input variables. Commun Stat: Simul Comput 1982;B11(3):311–34.
[152] Steinberg HA. Generalized quota sampling. Nuclear Sci Engng 1963;15:142–5.
[153] Raj D. Sampling theory. New York: McGraw-Hill; 1968.
[154] Stein M. Large sample properties of simulations using Latin hypercube sampling. Technometrics 1987;29(2):143–51.
[155] Owen AB. A central limit theorem for Latin hypercube sampling. J R Stat Soc. Ser B, Methodol. 1992;54(2):541–51.
[156] Billingsley P. Convergence of probability measures. New York: Wiley; 1968.
[157] Mathé P. Hilbert space analysis of Latin hypercube sampling. Proc Am Math Soc 2001;129(5):1477–92.
[158] Hoshino N, Takemura A. On reduction of finite-sample variance by extended Latin hypercube sampling. Bernoulli 2000;6(6): 1035–50.
[159] US NRC (US Nuclear Regulatory Commission). Reactor safety study—an assessment of accident risks in US commercial nuclear power plants. WASH-1400 (NUREG-75/014), Washington, DC: US Nuclear Regulatory Commission; 1975.
[160] Lewis HW, Budnitz RJ, Kouts HJC, Loewenstein WB, Rowe WD, von Hippel F, Zachariasen F. Risk assessment review group report to the US Nuclear Regulatory Commission. NUREG/CR-0400, Washington: US Nuclear Regulatory Commission; 1978.
[161] Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. SAND2001-0417, Albuquerque, NM: Sandia National Laboratories; 2002.
[162] McKay MD, Conover WJ, Whitehead DW. Report on the application of statistical techniques to the analysis of computer codes. Los Alamos, NM: Los Alamos Scientific Laboratory; 1976.
[163] Steck GP, Iman RL, Dahlgren DA. Probabilistic analysis of LOCA, Annual report for 1976. SAND76-0535, Albuquerque, NM: Sandia National Laboratories; 1976.
[164] Iman RL, Davenport JM, Ziegler DK. Latin hypercube sampling (program user's guide). SAND79-1473, Albuquerque, NM: Sandia National Laboratories; 1980.
[165] Iman RL, Shortencarier MJ. A FORTRAN 77 program and user's guide for the generation of Latin hypercube and random samples for use with computer models. NUREG/CR-3624, SAND83-2365, Albuquerque, NM: Sandia National Laboratories; 1984.
[166] Iman RL, Helton JC, Campbell JE. Risk methodology for geologic disposal of radioactive waste: sensitivity analysis techniques. SAND78-0912, NUREG/CR-0390, Albuquerque, NM: Sandia National Laboratories; 1978.
[167] Campbell JE, Dillon RT, Tierney MS, Davis HT, McGrath PE, Pearson FJ, Shaw HR, Helton JC, Donath FA. Risk methodology for geologic disposal of radioactive waste: interim report. SAND78-0029, NUREG/CR-0458, Albuquerque, NM: Sandia National Laboratories; 1978.
[168] Cranwell RM, Campbell JE, Helton JC, Iman RL, Longsine DE, Ortiz NR, Runkle GE, Shortencarier MJ. Risk methodology for geologic disposal of radioactive waste: final report. SAND81-2573, NUREG/CR-2452, Albuquerque, NM: Sandia National Laboratories; 1987.
[169] Sprung JL, Aldrich DC, Alpert DJ, Cunningham MA, Weigand GG. Overview of the MELCOR risk code development program. Proceedings of the International Meeting on Light Water Reactor Severe Accident Evaluation, Cambridge, MA, Boston, MA: Stone and Webster Engineering Corporation; 1983. pp. TS-10.1-1 to TS-10.1-8.
[170] Iman RL, Helton JC. A comparison of uncertainty and sensitivity analysis techniques for computer models. NUREG/CR-3904, SAND84-1461, Albuquerque, NM: Sandia National Laboratories; 1985.
[171] Iman RL, Helton JC. An investigation of uncertainty and sensitivity analysis techniques for computer models. Risk Anal 1988;8(1): 71–90.
[172] US NRC (US Nuclear Regulatory Commission). Severe accident risks: an assessment for five US nuclear power plants. NUREG-1150, vols. 1–3. Washington, DC: US Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Division of Systems Research; 1990–1991.
[173] Breeding RJ, Helton JC, Gorham ED, Harper FT. Summary description of the methods used in the probabilistic risk assessments for NUREG-1150. Nuclear Engng Des 1992;135(1):1–27.
[174] Breeding RJ, Helton JC, Murfin WB, Smith LN. Evaluation severe accident risks: Surry Unit 1. NUREG/CR-4551, SAND86-1309, vol. 3, Rev. 1. Albuquerque, NM: Sandia National Laboratories; 1990.
[175] Payne AC, Breeding RJ, Jow H-N, Helton JC, Smith LN, Shiver AW. Evaluation of severe accident risks: Peach Bottom Unit 2.

NUREG/CR-4551, SAND86-1309, vol. 4, Rev. 1. Albuquerque, NM: Sandia National Laboratories; 1990.
[176] Gregory JJ, Breeding RJ, Murfin WB, Helton JC, Higgins SJ, Shiver AW. Evaluation of severe accident risks: Sequoyah Unit 1. NUREG/CR-4551, SAND86-1309, vol. 5, Rev. 1. Albuquerque, NM: Sandia National Laboratories; 1990.
[177] Brown TD, Breeding RJ, Jow H-N, Helton JC, Higgins SJ, Amos CN, Shiver AW. Evaluation of severe accident risks: Grand Gulf Unit 1. NUREG/CR-4551, SAND86-1309, vol. 6, Rev. 1. Albuquerque, NM: Sandia National Laboratories; 1990.
[178] Park CK, Cazzoli EG, Grimshaw CA, Tingle A, Lee M, Pratt WT. Evaluation of severe accident risks: Zion Unit 1. NUREG/CR-4551, BNL/NUREG-5209, vol. 7, Rev. 1. Upton, NY: Brookhaven National Laboratory; 1993.
[179] Breeding RJ, Helton JC, Murfin WB, Smith LN, Johnson JD, Jow H-N, Shiver AW. The NUREG-1150 probabilistic risk assessment for the Surry Nuclear Power Station. Nuclear Engng Des 1992;135(1):29–59.
[180] Payne Jr. AC, Breeding RJ, Helton JC, Smith LN, Johnson JD, Jow H-N, Shiver AW. The NUREG-1150 probabilistic risk assessment for the Peach Bottom Atomic Power Station. Nuclear Engng Des 1992;135(1):61–94.
[181] Gregory JJ, Breeding RJ, Helton JC, Murfin WB, Higgins SJ, Shiver AW. The NUREG-1150 probabilistic risk assessment for the Sequoyah Nuclear Plant. Nuclear Engng Des 1992;135(1):92–115.
[182] Brown TD, Breeding RJ, Helton JC, Jow H-N, Higgins SJ, Shiver AW. The NUREG-1150 probabilistic risk assessment for the Grand Gulf Nuclear Station. Nuclear Engng Des 1992;135(1):117–37.
[183] Helton JC, Breeding RJ. Calculation of reactor accident safety goals. Reliab Engng Syst Saf 1993;39(2):129–58.
[184] Payne Jr.AC. Analysis of the LaSalle Unit 2 Nuclear Power Plant: risk methods integration and evaluation program (RMIEP). Summary. NUREG/CR-4832; SAND92-0537, vol. 1. Albuquerque, NM: Sandia National Laboratories; 1992.
[185] Payne Jr.AC, Sype TT, Whitehead DW, Shiver AW. Analysis of the LaSalle Unit 2 Nuclear Power Plant: risk methods integration and evaluation program (RMIEP). SAND92-0537, vol. 2. Albuquerque, NM: Sandia National Laboratories; 1992.
[186] Payne Jr.AC, Daniel SL, Whitehead DW, Sype TT, Dingman SE, Shaffer CJ. Analysis of the LaSalle Unit 2 Nuclear Power Plant: risk methods integration and evaluation program (RMIEP) internal events accident sequence qualification. Main report. SAND92-0537, vol. 3, pt. 1. Albuquerque, NM: Sandia National Laboratories; 1992.
[187] Payne Jr.AC, Daniel SL, Whitehead DW, Sype TT, Dingman SE, Shaffer CJ. Analysis of the LaSalle Unit 2 Nuclear Power Plant: risk methods integration and evaluation program (RMIEP) internal events accident sequence quantification, appendices. SAND92-0537, vol. 3, pt. 2. Albuquerque, NM: Sandia National Laboratories; 1992.
[188] NAS/NRC (National Academy of Sciences/National Research Council). The Waste Isolation Pilot Plant, a potential solution for the disposal of transuranic waste. Washington DC: National Academy Press, Committee on the Waste Isolation Pilot Plant, Board on Radioactive Waste Management, Commission on Geosciences, Environment, and Resources, National Research Council. 1996.
[189] Rechard RP. Historical background on performance assessment for the Waste Isolation Pilot Plant. Reliab Engng Syst Saf 2000;69(1–3):5–46.
[190] US DOE (US Department of Energy). Title 40 CFR Part 191 compliance certification application for the Waste Isolation Pilot Plant, DOE/CAO-1996-(2184). Carlsbad, NM: US Department of Energy, Carlsbad Area Office. Waste Isolation Pilot Plant. 1996.
[191] Helton JC, Marietta MG. Special issue: the 1996 performance assessment for the Waste Isolation Pilot Plant. Reliab Engng Syst Saf 2000;69(1–3):1–451.
[192] US EPA (US Environmental Protection Agency). 40 CFR Part 191: environmental standards for the management and disposal of spent nuclear fuel, high-level and transuranic radioactive wastes; final rule. Fed Register 1985;50(182):38066–89.
[193] US EPA (US Environmental Protection Agency). 40 CFR Part 194: criteria for the certification and re-certification of the Waste Isolation Pilot Plant's compliance with the 40 CFR Part 191 disposal regulations; final rule. Fed Register 1996;61(28):5224–45.
[194] Helton JC. Risk, uncertainty in risk, and the EPA release limits for radioactive waste disposal. Nuclear Technol 1993;101(1):18–39.
[195] Helton JC, Anderson DR, Marietta MG, Rechard RP. Performance assessment for the Waste Isolation Pilot Plant: from regulation to calculation for 40 CFR 191.13. Oper Res 1997;45(2):157–77.
[196] US DOE (US Department of Energy). Viability assessment of a repository at Yucca Mountain. DOE/RW-0508, Washington, DC: US Department of Energy, Office of Civilian Radioactive Waste Management; 1998.
[197] CRWMS M&O (Civilian Radioactive Waste Management System Management and Operating Contractor). Total system performance assessment for the site recommendation. TDR-WIS-PA-000001 REV 00, Las Vegas, NV: CRWMS M&O; 2000.
[198] CRWMS M&O (Civilian Radioactive Waste Management System Management and Operating Contractor). Total system performance assessment (TSPA) model for site recommendation. MDL-WIS-PA-000002 REV 00, Las Vegas, NV: CRWMS M&O; 2000.
[199] Kincaid CT, Eslinger PW, Nichols WE, Bunn AL, Bryce RW, Miley TB, Richmond MC, Snyder SF, Aaberg RL. Groundwater/vadose zone integration project: system assessment capacity (Revision 0) assessment description, requirements, software design, and test plan, BHI-01365. Richland, WA: Bechtel Hanford, Inc; 2000.
[200] BHI (Bechtel Hanford, Inc). Groundwater/vadose zone integration project: preliminary system assessment capability concepts for architecture, platform and data management. CCN 0512242, Richland, WA: BHI; 1999.
[201] Gilbert RO, Bittner EA, Essington EH. On the use of uncertainty analyses to test hypotheses regarding deterministic model predictions of environmental processes. J Environ Radioact 1995;27(3):231–60.
[202] Hyman TC, Hamby DM. Parameter uncertainty and sensitivity in a liquid-effluent dose model. Environ Monitor Assess 1995;38(1):51–65.
[203] Toran L, Sjoreen A, Morris M. Sensitivity analysis of solute transport in fractured porous media. Geophys Res Lett 1995;22(11):1433–6.
[204] Anderson RM, Donnelly CA, Ferguson NM, Woolhouse MEJ, Watt CJ, Udy HJ, MaWhinney S, Dunstan SP, Southwood TRE, Wilesmith JW, Ryan JBM, Hoinville LJ, Hillerton JE, Austin AR, Wells GAH. Transmission dynamics and epidemiology of BSE in British cattle. Nature 1996;382(6594):779–88.
[205] Fish DJ, Burton MR. The effect of uncertainties in kinetic and photochemical data on model predictions of stratosphere ozone depletion. J Geophys Res 1997;102(D21):25537–42.
[206] Keramat M, Kielbasa R. Latin hypercube sampling Monte Carlo estimation of average quality index for integrated circuits. Analog Integrated Circuits Signal Process 1997;14(1–2):131–42.
[207] Ellerbroek DA, Durnford DS, Loftis JC. Modeling pesticide transport in an irrigated field with variable water application and hydraulic conductivity. J Environ Qual 1998;27(3):495–504.
[208] Chen C, Finch SJ, Mendell NR, Gordon D. Comparison of empirical strategies to maximize GENEHUNTER Lod scores. Genet Epidemiol 1999;17(S1):S103–8.
[209] Considine DB, Stolarski RS, Hollandsworth SM, Jackman CH, Fleming EL. A Monte Carlo uncertainty analysis of ozone trend predictions in a two-dimensional model. J Geophys Res 1999;104(D1):1749–65.
[210] Kros J, Pebesma EJ, Reinds GJ, Finke PA. Uncertainty assessment in modeling soil acidification at the European scale: a case study. J Environ Qual 1999;28(2):366–77.

[211] Mrawira D, Welch WJ, Schonlau M, Haas R. Sensitivity analysis of computer models: World Bank HDM-III model. J Transport Engng 1999;125(5):421–8.
[212] Padmanabhan SK, Pitchumani R. Stochastic modeling of non-isothermal flow during resin transfer molding. Int J Heat Mass Transfer 1999;42(16):3057–70.
[213] Soutter M, Musy A. Global sensitivity analyses of three pesticide leaching models using a Monte Carlo approach. J Environ Qual 1999;28(4):1290–7.
[214] Fischer F, Hasemann I, Jones JA. Techniques applied in the COSYMA accident consequence uncertainty analysis. Radiat Prot Dosimetry 2000;90(3):317–23.
[215] Oh BH, Yang IH. Sensitivity analysis of time-dependent behavior in PSC box girder bridges. J Struct Engng 2000;126(2):171–9.
[216] Morris MD. Three technometrics experimental design classics. Technometrics 2000;42(1):26–7.
[217] Helton JC. Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. Reliab Engng Syst Saf 1993;42(2–3):327–67.
[218] Conover WJ. Practical nonparametric statistics, 2nd ed. New York: Wiley; 1980.
[219] Golub GH, van Loan CF. Matrix computations. Baltimore: Johns Hopkins University Press; 1983.
[220] Anderson TW. An introduction to multivariate statistical analysis, 2nd ed. New York: Wiley; 1984.
[221] Iman RL, Davenport JM. Rank correlation plots for use with correlated input variables in simulation studies, SAND80-1903. Albuquerque, NM: Sandia National Laboratories; 1980.
[222] Iman RL, Davenport JM. Rank correlation plots for use with correlated input variables. Commun Stat: Simul Comput 1982; B11(3):335–60.
[223] Ye KQ, Li W, Sudjianto A. Algorithmic construction of optimal symmetric Latin hypercube designs. J Stat Plann Inference 2000; 90(1):145–59.
[224] Pebesma EJ, Heuvelink GBM. Latin hypercube sampling of gaussian random fields. Technometrics 1999;41(4):303–12.
[225] Avramidis AN, Wilson JR. Correlation-induction techniques for estimating quantiles in simulation experiments. Oper Res 1998; 46(4):574–91.
[226] Ye KQ. Orthogonal column Latin hypercubes and their application in computer experiments. J Am Stat Assoc 1998;93(444):1430–9.
[227] Tang BX. Selecting Latin hypercubes using correlation criteria. Stat Sin 1998;8(3):965–77.
[228] Salagame RR, Barton RR. Factorial hypercube designs for spatial correlation regression. J Appl Stat 1997;24(4):453–73.
[229] Owen AB. Monte Carlo variance of scrambled net quadrature. Siam J Numer Anal 1997;34(5):1884–910.
[230] Loh W-L. Estimating the integral of a squared regression function with Latin hypercube. Stat Probab Lett 1997;31(4):339–49.
[231] Loh W-L. On Latin hypercube sampling. Ann Stat 1996;24(5): 2058–80.
[232] Avramidis AN, Wilson JR. Integrated variance reduction strategies for simulation. Oper Res 1996;44(2):327–46.
[233] Loh W-L. A combinatorial central limit theorem for randomized orthogonal array sampling designs. Ann Stat 1996;24(3):1209–24.
[234] Streltsov S, Vakili P. Variance reduction algorithms for parallel replicated simulation of uniformized Markov chains. Discrete Event Dyn Syst: Theor Appl 1996;6(2):159–80.
[235] Harris CM, Hoffman KL, Yarrow L-A. Obtaining minimum-correlation Latin hypercube sampling plans using an IP-based heuristic. OR Spektrum 1995;17(2–3):139–48.
[236] Owen AB. Controlling correlations in Latin hypercube samples. J Am Stat Assoc 1994;89(428):1517–22.
[237] Owen A. Lattice sampling revisited: Monte-Carlo variance of means over randomized orthogonal arrays. Ann Stat 1994;22(2): 930–45.
[238] Tang BX. A theorem for selecting OA-based Latin hypercubes using a distance criterion. Commun Stat—Theor Meth 1994;23(7): 2047–58.
[239] Park J-S. Optimal Latin-hypercube designs for computer experiments. J Stat Plann Inference 1994;39(1):95–111.
[240] Tang BX. Orthogonal array-based Latin hypercubes. J Am Stat Assoc 1993;88(424):1392–7.
[241] Owen AB. Orhogonal arrays for computer experiments, integration and visualization. Stat Sin 1992;2(2):439–52.
[242] Iman RL, Conover WJ. Small sample sensitivity analysis techniques for computer models, with an application to risk assessment. Commun Stat: Theor Meth 1980;A9(17):1749–842.
[243] Beckman RJ, McKay MD. Monte-Carlo estimation under different distribution using the same simulation. Technometrics 1987;29(2): 153–60.
[244] Iman RL. Statistical methods for including uncertainties associated with the geologic isolation of radioactive waste which allow for a comparison with licensing criteria. In: Kocher DC, editor. Proceedings of the Symposium on Uncertainties Associated with the Regulation of the Geologic Disposal of High-Level Radioactive Waste, Gatlinburg, TN. NUREG/CP-0022; CONF-810372, Washington, DC: US Nuclear Regulatory Commission, Directorate of Technical Information and Document Control; 1981. p. 145–57.
[245] Helton JC, Bean JE, Berglund JW, Davis FJ, Economy K, Garner JW, Johnson JD, MacKinnon RJ, Miller J, O'Brien DG, Ramsey JL, Schreiber JD, Shinta A, Smith LN, Stoelzel DM, Stockman C, Vaughn P. Uncertainty and sensitivity analysis results obtained in the 1996 performance assessment for the Waste Isolation Pilot Plant. SAND98-0365, Albuquerque, NM: Sandia National Laboratories; 1998.
[246] Vaughn P, Bean JE, Helton JC, Lord ME, MacKinnon RJ, Schreiber JD. Representation of two-phase flow in the vicinity of the repository in the 1996 performance assessment for the Waste Isolation Pilot Plant. Reliab Engng Syst Saf 2000;69(1–3):205–26.
[247] Helton JC, Bean JE, Economy K, Garner JW, MacKinnon RJ, Miller J, Schreiber JD, Vaughn P. Uncertainty and sensitivity analysis for two-phase flow in the vicinity of the repository in the 1996 performance assessment for the Waste Isolation Pilot Plant: undisturbed conditions. Reliab Engng Syst Saf 2000;69(1–3): 227–61.
[248] Helton JC, Bean JE, Economy K, Garner JW, MacKinnon RJ, Miller J, Schreiber JD, Vaughn P. Uncertainty and sensitivity analysis for two-phase flow in the vicinity of the repository in the 1996 performance assessment for the Waste Isolation Pilot Plant: disturbed conditions. Reliab Engng Syst Saf 2000;69(1–3):263–304.
[249] Helton JC, Anderson DR, Jow H-N, Marietta MG, Basabilvazo G. Performance assessment in support of the 1996 compliance certification application for the Waste Isolation Pilot Plant. Risk Anal 1999;19(5):959–86.
[250] Helton JC, Martell M-A, Tierney MS. Characterization of subjective uncertainty in the 1996 performance assessment for the Waste Isolation Pilot Plant. Reliab Engng Syst Saf 2000;69(1–3):191–204.
[251] Ibrekk H, Morgan MG. Graphical communication of uncertain quantities to nontechnical people. Risk Anal 1987;7(4):519–29.
[252] Iman RL, Helton JC. The repeatability of uncertainty and sensitivity analyses for complex probabilistic risk assessments. Risk Anal 1991; 11(4):591–606.
[253] Helton JC, Johnson JD, McKay MD, Shiver AW, Sprung JL. Robustness of an uncertainty and sensitivity analysis of early exposure results with the MACCS reactor accident consequence model. Reliab Engng Syst Saf 1995;48(2):129–48.
[254] Iman RL, Conover WJ. The use of the rank transform in regression. Technometrics 1979;21(4):499–509.
[255] Kleijnen JPC, Helton JC. Statistical analyses of scatterplots to identify important factors in large-scale simulations, 1: review and comparison of techniques. Reliab Engng Syst Saf 1999;65(2): 147–85.

[256] Kleijnen JPC, Helton JC. Statistical analyses of scatterplots to identify important factors in large-scale simulations, 2: robustness of techniques. Reliab Engng Syst Saf 1999;65(2):187–97.
[257] Iman RL, Conover WJ. A measure of top-down correlation. Technometrics 1987;29(3):351–7.
[258] Assunçao R. Testing spatial randomness by means of angles. Biometrics 1994;50:531–7.
[259] Ripley BD. Spatial point pattern analysis in ecology. In: Legendre P, Legendre L, editors. Developments in numerical ecology. NATO ASI series, series G: ecological sciences, vol. 14. Berlin: Springer; 1987. p. 407–30.
[260] Zeng G, Dubes RC. A comparison of tests for randomness. Pattern Recog 1985;18(2):191–8.
[261] Diggle PJ, Cox TF. Some distance-based tests of independence for sparsely-sampled multivariate spatial point patterns. Int Stat Rev 1983;51(1):11–23.
[262] Byth K. On robust distance-based intensity estimators. Biometrics 1982;38(1):127–35.
[263] Byth K, Ripley BD. On sampling spatial patterns by distance methods. Biometrics 1980;36(2):279–84.
[264] Diggle PJ. On parameter estimation and goodness-of-fit testing for spatial point patterns. Biometrics 1979;35(1):87–101.
[265] Diggle PJ. Statistical methods for spatial point patterns in ecology. In: Cormack RM, Ord JK, editors. Spatial and temporal analysis in ecology. Fairfield, MD: International Co-operative Publishing House; 1979. p. 95–150.
[266] Besag J, Diggle PJ. Simple Monte Carlo tests for spatial pattern. Appl Stat 1977;26(3):327–33.
[267] Diggle PJ, Besag J, Gleaves JT. Statistical analysis of spatial point patterns by means of distance methods. Biometrics 1976;32:659–67.
[268] Ripley BD. Tests of randomness for spatial point patterns. J R Stat Soc 1979;41(3):368–74.
[269] Cox TF, Lewis T. A conditioned distance ratio method for analyzing spatial patterns. Biometrika 1976;63(3):483–91.
[270] Holgate P. The use of distance methods for the analysis of spatial distribution of points. In: Lewis PAW, editor. Stochastic point processes: statistical analysis, theory, and applications. New York: Wiley; 1972. p. 122–35.
[271] Holgate P. Tests of randomness based on distance methods. Biometrika 1965;52(3–4):345–53.
[272] Garvey JE, Marschall EA, Wright RA. From star charts to stoneflies: detecting relationships in continuous bivariate data. Ecology 1998; 79(2):442–7.
[273] Fasano G, Franceschini A. A multidimensional version of the Kolmogorov–Smirnov test. Mon Not R Astron Soc 1987;225(1): 155–70.
[274] Gosset E. A 3-dimensional extended Kolmogorov–Smirnov test as a useful tool in astronomy. Astron Astrophys 1987;188(1):258–64.
[275] Peacock JA. Two-dimensional goodness-of-fit testing in astronomy. Mon Not R Astron Soc 1983;202(2):615–27.
[276] Helton JC, Davis FJ. Sampling-based methods for uncertainty and sensitivity analysis. SAND99-2240, Albuquerque, NM: Sandia National Laboratories; 2000.
[277] Hamby DM. A comparison of sensitivity analysis techniques. Health Phys 1995;68(2):195–204.
[278] Hamby DM. A review of techniques for parameter sensitivity analysis of environmental models. Environ Monitor Assess 1994; 32(2):135–54.
[279] Saltelli A, Andres TH, Homma T. Sensitivity analysis of model output. An investigation of new techniques. Comput Stat Data Anal 1993;15(2):211–38.
[280] Saltelli A, Marivoet J. Non-parametric statistics in sensitivity analysis for model output. A comparison of selected techniques. Reliab Engng Syst Saf 1990;28(2):229–53.
[281] Iman RL, Helton JC, Campbell JE. An approach to sensitivity analysis of computer models, part 1. Introduction, input variable selection and preliminary variable assessment. J Qual Technol 1981; 13(3):174–83.
[282] Iman RL, Helton JC, Campbell JE. An approach to sensitivity analysis of computer models, part 2. Ranking of input variables, response surface validation, distribution effect and technique synopsis. J Qual Technol 1981;13(4):232–40.
[283] Kaplan S, Garrick BJ. On the quantitative definition of risk. Risk Anal 1981;1(1):11–27.
[284] Parry GW, Winter PW. Characterization and evaluation of uncertainty in probabilistic risk analysis. Nuclear Saf 1981;22(1): 28–42.
[285] Haan CT. Parametric uncertainty in hydrologic modeling. Trans ASAE 1989;32(1):137–46.
[286] Apostolakis G. The concept of probability in safety assessments of technological systems. Science 1990;250(4986):1359–64.
[287] Helton JC. Treatment of uncertainty in performance assessments for complex systems. Risk Anal 1994;14(4):483–511.
[288] Hoffman FO, Hammonds JS. Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. Risk Anal 1994; 14(5):707–12.
[289] Paté-Cornell ME. Uncertainties in risk analysis: six levels of treatment. Reliab Engng Syst Saf 1996;54(2–3):95–111.
[290] Winkler RL. Uncertainty in probabilistic risk assessment. Reliab Engng Syst Saf 1996;54(2–3):127–32.
[291] Barnett V, O'Hagan A. Setting environmental standards: the statistical approach to handling uncertainty and variation. London: Chapman & Hall; 1997.
[292] Cullen A. Addressing uncertainty-lessons from probabilistic exposure analysis. Inhal Toxicol 1999;11(6–7):603–10.
[293] Francis RICC, Shotton R. Risk in fisheries management: a review. Can J Fish Aquat Sci 1997;54(8):1699–715.
[294] Cullen AC, Frey HC. Probabilistic techniques in exposure assessment: a handbook for dealing with variability and uncertainty in models and inputs. London: Plenum Press; 1999.
[295] Kelly EJ, Campbell K. Separating variability and uncertainty in environmental risk assessment-making choices. Hum Ecol Risk Assess 2000;6(1):1–13.
[296] Hacking I. The emergence of probability: a philosophical study of early ideas about probability, induction and statistical inference. London: Cambridge University Press; 1975.
[297] Shafer G. Non-additive probabilities in work of Bernoulli and Lambert. Arch Hist Exact Sci 1978;19(4):309–70.
[298] Bernstein PL. Against the Gods: the remarkable story of risk. New York: Wiley; 1996.
[299] Lohman K, Pai P, Seigneur C, Mitchell D, Heim K, Wandland K, Levin L. A probabilistic analysis of regional mercury impacts on wildlife. Hum Ecol Risk Assess 2000;6(1):103–30.
[300] Maxwell RM, Kastenberg WE. Stochastic environmental risk analysis: an integrated methodology for predicting cancer risk from contaminated groundwater. Stochast Environ Res Risk Assess 1999;13(1–2):27–47.
[301] Maxwell RM, Kastenberg WE. A model for assessing and managing the risks of environmental lead emissions. Stochast Environ Res Risk Assess 1999;13(4):231–50.
[302] Maxwell RM, Kastenberg WE, Rubin Y. A methodology to integrate site characterization information into groundwater-driven health risk assessment. Water Resour Res 1999;35(9):2841–55.
[303] McKone TE. Uncertainty and variability in human exposures to soil contaminants through home-grown food: a Monte Carlo assessment. Risk Anal 1994;14(4):449–63.
[304] Allen BC, Covington TR, Clewell HJ. Investigation of the impact of pharmacokinetic variability and uncertainty on risks predicted with a pharmacokinetic model for chloroform. Toxicology 1996;111(1–3): 289–303.
[305] Øvreberg O, Damsleth E, Haldorsen HH. Putting error bars on reservoir engineering forecasts. J Petrol Technol 1992;44(6):732–8.

[306] Price PS, Su SH, Harrington JR, Keenan RE. Uncertainty and variation of indirect exposure assessments: an analysis of exposure to tetrachlorodibenzene-*p*-dioxin from a beef consumption pathway. Risk Anal 1996;16(2):263–77.

[307] PLG (Pickard, Lowe and Garrick, Inc). Seabrook Station Preliminary Probabilistic Safety Assessment, PLG-0300, vols. 1–6, Summary. Prepared for Public Service Company of New Hampshire, Manchester and Yankee Atomic Electric Company, Framingham, MA. Irvine, CA: Pickard, Lowe and Garrick, Inc. 1983

[308] PLG (Pickard, Lowe and Garrick, Inc., Westinghouse Electric Corporation, and Fauske & Associates, Inc). Indian Point Probabilistic Safety Study, Prepared for the Power Authority of the State of New York and Consolidated Edison Company of New York, Inc. Irvine, CA: Pickard, Lowe and Garrick, Inc. 1982

[309] Howard BA, Crawford MB, Galson DA, Marietta MG. Regulatory basis for the Waste Isolation Pilot Plant performance assessment. Reliab Engng Syst Saf 2000;69(1–3):109–27.

[310] Farmer FR. Reactor safety and siting: a proposed risk criterion. Nuclear Saf 1967;8(6):539–48.

[311] Cox DC, Baybutt P. Limit lines for risk. Nuclear Technol 1982; 57(3):320–30.

[312] Munera HA, Yadigaroglu G. On farmer's line, probability density functions, and overall risk. Nuclear Technol 1986;74(2):229–32.

[313] Helton JC, Johnson JD, Jow H-N, McCurley RD, Rahal LJ. Stochastic and subjective uncertainty in the assessment of radiation exposure at the Waste Isolation Pilot Plant. Hum Ecol Risk Assess 1998;4(2):469–526.

[314] Galson DA, Swift PN, Anderson DR, Bennett DG, Crawford MB, Hicks TW, Wilmot RD, Basabilvazo G. Scenario development for the Waste Isolation Pilot Plant compliance certification application. Reliab Engng Syst Saf 2000;69(1–3):129–49.

[315] Helton JC, Davis FJ, Johnson JD. Characterization of stochastic uncertainty in the 1996 performance assessment for the Waste Isolation Pilot Plant. Reliab Engng Syst Saf 2000;69(1–3): 167–89.

[316] Stoelzel DM, O'Brien DG, Garner JW, Helton JC, Johnson JD, Smith LN. Direct releases to the surface and associated complementary cumulative distribution functions in the 1996 performance assessment for the Waste Isolation Pilot Plant: direct brine release. Reliab Engng Syst Saf 2000;69(1–3):343–67.

[317] Berglund JW, Garner JW, Helton JC, Johnson JD, Smith LN. Direct releases to the surface and associated complementary cumulative distribution functions in the 1996 performance assessment for the Waste Isolation Pilot Plant: cuttings, cavings and spallings. Reliab Engng Syst Saf 2000;69(1–3):305–30.

[318] LaVenue AM. Analysis of the generation of transmissivity fields for the Culebra dolomite. Albuquerque, NM: Sandia National Laboratories; 1996. Available in Sandia WIPP Records Center as WPO#40517.

[319] LaVenue AM, RamaRao BS. A modeling approach to address spatial variability within the Culebra dolomite transmissivity field. SAND92-7306, Albuquerque, NM: Sandia National Laboratories; 1992.

[320] Stockman CT, Garner JW, Helton JC, Johnson JD, Shinta A, Smith LN. Radionuclide transport in the vicinity of the repository and associated complementary cumulative distribution functions in the 1996 performance assessment for the Waste Isolation Pilot Plant. Reliab Engng Syst Saf 2000;69(1–3):397–420.

[321] Stone CM. SANTOS—a two-dimensional finite element program for the quasistatic, large deformation, inelastic response of solids. SAND90-0543, Albuquerque, NM: Sandia National Laboratories; 1997.

[322] Stone CM. Final disposal room structural response calculations. SAND97-0795, Albuquerque, NM: Sandia National Laboratories; 1997.

[323] Ramsey JL, Blaine R, Garner JW, Helton JC, Johnson JD, Smith LN, Wallace M. Radionuclide and colloid transport in the Culebra dolomite and associated complementary cumulative distribution functions in the 1996 performance assessment for the Waste Isolation Pilot Plant. Reliab Engng Syst Saf 2000;69(1–3):397–420.

[324] Helton JC, Anderson DR, Basabilvazo G, Jow H-N, Marietta MG. Summary discussion of the 1996 performance assessment for the Waste Isolation Pilot Plant. Reliab Engng Syst Saf 2000;69(1–3): 437–51.

[325] Helton JC. Uncertainty and sensitivity analysis in performance assessment for the Waste Isolation Pilot Plant. Comput Phys Commun 1999;117(1–2):156–80.